

**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN**  
**FACULTAD DE PSICOLOGÍA**



**“ESTUDIO DE DATOS NORMATIVOS DE TRES  
PRUEBAS NEUROPSICOLÓGICAS EN POBLACIÓN  
ADULTA DE MEXICALI, BAJA CALIFORNIA”**

**PRESENTA:**

**FRANCISCO JAVIER GALARZA DEL ÁNGEL**

**Tesis como requisito parcial para obtener el grado de:**

**DOCTORADO EN FILOSOFÍA  
CON ESPECIALIDAD EN PSICOLOGÍA**

**MONTERREY, N. L., MARZO DE 2021**

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN



**ESTUDIO DE DATOS NORMATIVOS DE TRES PRUEBAS  
NEUROPSICOLÓGICAS EN POBLACIÓN ADULTA DE  
MEXICALI, BAJA CALIFORNIA**

TESIS:

QUE PARA OBTENER EL GRADO DE DOCTOR EN  
FILOSOFIA CON ESPECIALIDAD EN PSICOLOGIA  
PRESENTA

Francisco Javier Galarza del Ángel

Director de Tesis: Dr. Arnoldo Téllez López

Co-director: Dr. Diego Fernando Rivera Camacho

Monterrey, Nuevo León, Marzo de 2021

**UNIVERSIDAD AUTONOMA DE NUEVO LEON**  
**FACULTAD DE PSICOLOGIA**  
**SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO**

**DOCTORADO EN FILOSOFÍA CON ORIENTACIÓN EN PSICOLOGÍA**

La presente tesis titulada "Estudio de datos normativos de tres pruebas neuropsicológicas en población adulta de Mexicali, Baja California" ha sido aprobada por el comité de tesis.



---

Dr. Arnoldo Téllez López  
Director de tesis



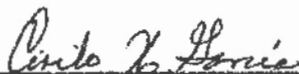
---

Dr. Diego Fernando Rivera Camacho  
Codirector de tesis



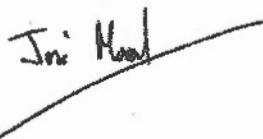
---

Dra. Xóchitl Angélica Ortiz Jiménez  
Revisor de tesis



---

Dr. Cirilo Humberto García Cadena  
Revisor de tesis



---

Dr. José Moral de la Rubia  
Revisor de tesis

Monterrey, N. L., México, diciembre de 2020

## Agradecimientos

Todo trabajo de investigación necesita la colaboración y estímulo de muchas personas que hacen posible el desarrollo del trabajo. Son horas de planeación, diseño de estrategias, análisis de datos, elaboración de análisis estadísticos, tablas, cuadros, etcetera. Este trabajo contó no es la excepción.

En primer lugar agradezco la colaboración del equipo de estudiantes colaboradores del Laboratorio de Psicofisiología, de la Facultad de Ciencias Humanas de la Universidad Autónoma de Baja California, quienes fueron los encargados de realizar las evaluaciones de las personas que participaron y de quienes obtuvimos los datos.

Mi agradecimiento para las personas que accedieron a ser evaluados ya que sin su ayuda esta tesis no hubiera podido desarrollarse.

El Dr. Víctor Manuel Padilla Montemayor siempre apoyó este trabajo y siempre estuvo dispuesto a colaborar en las cuestiones académicas y administrativas.

Debo mencionar que estoy en deuda con el Dr. Arnoldo Téllez López, quien me apoyó desde el inicio del Doctorado y quién siempre estuvo al pendiente de los obstáculos que se iban presentando en esta travesía. El Dr. Arnoldo Téllez, director de esta tesis, siempre me empujaba con firmeza, y tal vez con algunas directrices hipnóticas, a que no me diera por vencido.

Un agradecimiento especial al Mtro. Luis Alfredo Padilla López. colega y

amigo desde hace muchos años, quien también me apoyaba...a su estilo, pero "me picaba la cresta" y con sus comentarios y discusiones a este trabajo me sugería modificaciones a la tesis.

No puedo dejar de mencionar el gran apoyo que he recibido del Dr. Juan Carlos Arango Lasprilla, quien me invitó a colaborar en el Estudio Multicéntrico del que se desprendieron los datos para esta tesis. En este estudio colaboramos equipos de investigación de 12 países de Latinoamérica y pudimos evaluar a casi doce mil personas. El Dr. Arango Lasprilla revisó este trabajo desde el inicio y gracias a su apoyo pude realizar dos estancias en su laboratorio para revisar los datos y escribir gran parte de esta tesis.

Otra persona quien me apoyó desde que nos conocimos es el Dr. Diego Fernando Rivera Camacho, Co-director de esta tesis. El Dr. Rivera Camacho es Profesor en el Departamento de Salud de la Universidad Pública de Navarra, y pasó largas jornadas ayudándome en la elaboración de los análisis estadísticos de esta tesis. Estaré en deuda por siempre.

A mi familia en Monterrey, N. L. por todo el apoyo que me dieron, casa, alimentos, transporte y cariño. A mis hermanos y hermanas, cuñados y cuñada, sobrinos...a toda la familia muchas gracias.

Por último quiero agradecer y dedicarles esta tesis a mi esposa María Luisa y a mis hijos Andrea, Héctor y Aminda por su apoyo y comprensión en las jornadas en las que me ausentaba por las clases del doctorado y por las estancias que hice para continuar con la elaboración de esta tesis. Indudablemente, sin su aliento, motivación y apoyo yo no podría haber escrito ni una línea de esta tesis.

## Índice

RESUMEN .....	6
ABSTRACT .....	7
Capítulo 1. Marco teórico.....	8
1.1 Funciones ejecutivas. ....	8
1.2 Pruebas neuropsicológicas que miden las funciones ejecutivas.....	12
1.2.1 La Prueba de Colores y Palabras (Stroop).....	12
1.2.2 Prueba Modificada de Clasificación de Tarjetas de Wisconsin. ....	18
1.2.3 Prueba de modalidades símbolo-dígito (SDMT). ....	23
1.2.3.1 Estudios con datos normativos. ....	27
Capítulo 2. Planteamiento del problema .....	30
2.1. Planteamiento del problema .....	30
2.2 Objetivo general .....	32
2.3 Objetivos específicos .....	32
2.4 Datos demográficos de Baja California y Mexicali.....	33
2.5 Datos demográficos de Mexicali.....	34
2.6 Justificación .....	37
Capítulo 3. Método.....	39
3.1 Método.....	39
3.1.1 Participantes y tamaño de la muestra.....	39
3.1.2 Criterios de inclusión y exclusión. ....	39
3.2 Instrumentos .....	41
3.2.1 Entrevista inicial.....	41
3.2.2 Pruebas de screening. ....	42
3.2.3 Pruebas neuropsicológicas. ....	45
3.3 Procedimiento.....	47
3.3.1 Diseño de investigación utilizado. ....	47
3.3.2 Recolección de datos. ....	47
3.3.3. Análisis de datos.....	49
3.3.4 Aspectos Éticos .....	53

Capítulo 4. Resultados.....	54
4.1 Resultados.....	54
4.1.1 Análisis exploratorios de los datos.....	54
4.1.2 Efectos de las variables demográficas.....	68
4.1.3 Datos normativos para las puntuaciones estudiadas.....	77
4.1.4 Calculadora de datos normativos. ....	81
4.1.5 Comparaciones de modelos de regresión de 2015 y de 2020.....	82
Capítulo 5. Discusión .....	86
5.1 Discusión .....	86
5.2 Implicaciones.....	94
5.3 Limitaciones.....	96
Capítulo 6. Conclusión.....	98
Referencias .....	100

## RESUMEN

Las pruebas neuropsicológicas se utilizan en la evaluación y diagnóstico de los procesos cognoscitivos de pacientes. Sin embargo, la mayoría de estas pruebas no cuentan con datos normativos para México, lo que supone un problema al revisar la ejecución de las personas dado que se comparan contra baremos de otros países. Con el propósito de generar datos normativos para tres pruebas neuropsicológicas ampliamente utilizadas en México, se aplicaron la Prueba Modificada de Clasificación de Tarjetas de Wisconsin, la Prueba de Palabras y Colores de Stroop y la Prueba Símbolo-Dígito a 430 adultos de la ciudad de Mexicali, Baja California. Este trabajo forma parte de un Estudio multicéntrico en el que se evaluaron 5402 adultos en 20 ciudades de 12 países latinoamericanos.

Se utilizó un diseño correlacional explicativo y los participantes fueron asignados a diferentes grupos dependiendo de sus sexo, edad y escolaridad; una vez que cumplieran con los criterios de inclusión.

Análisis de datos: La muestra se dividió en dos sub-muestras de manera aleatoria. La primera estuvo formada por el 80% de los datos y la segunda por el 20% restante. La sub-muestra 1 se utilizó para generar los datos normativos, en tanto que la sub-muestra 2 se usó para comprobar los modelos generados en la sub-muestra 1. Los efectos de las variables demográficas se evaluaron mediante análisis de regresión múltiple que incluyó como variables predictoras la edad, edad<sup>2</sup>, escolaridad, escolaridad<sup>2</sup> y el sexo, así como las interacciones con estas variables. Las puntuaciones de la sub-muestra 2 fueron transformadas a puntuación  $z_i$  y a percentiles a partir de los datos normativos publicados en el año 2015 (Rivera y et al., 2015; Arango-Lasprilla y et al., 2015a y 2015b); es decir, cada puntuación directa de la sub-muestra 2 fue transformada a puntuaciones  $z_i$  y a percentil usando dos modelos: el primer modelo resultante de este trabajo y el segundo modelo publicado en el 2015.

Resultados: Se obtuvieron datos normativos para las tres pruebas neuropsicológicas y se organizaron en tablas de percentiles. Se observaron efectos lineales de la edad y de la escolaridad en el desempeño de los participantes. No se encontraron efectos del sexo sobre la ejecución. Adicionalmente se obtuvieron los modelos de regresión múltiple para cada prueba y sub-prueba. Al comparar los modelos 2015 y 2020 se encontró que el modelo 2020 explica un mayor porcentaje de varianza en todas las pruebas y subpruebas comparadas.

Conclusión: Se obtuvieron datos normativos sin sesgo de edad, escolaridad y sexo en las tres pruebas neuropsicológicas aplicadas. En todas las pruebas y sub-pruebas se demostró el efecto de la edad y de la escolaridad. El uso de técnicas de regresión múltiple produce datos normativos sin sesgo de edad, escolaridad y sexo.

Palabras clave: Datos normativos, pruebas neuropsicológicas, regresión múltiple, adultos.



## ABSTRACT

Neuropsychological tests are used in the evaluation and diagnosis of the cognitive processes of patients. However, most of these tests do not have normative data for Mexico, this situation is a problem at the moment of reviewing the execution of individuals since they are compared against normative data from other countries. In order to generate normative data for three neuropsychological tests widely used in Mexico; the Modified Wisconsin Card Sorting Test, the Word and Color Stroop Test, and the Symbol-Digit Modalities Test were applied to 430 adults from the city of Mexicali, Baja California. This work is part of a multicenter study in which 5,402 adults were evaluated in 20 cities of 12 Latin American countries.

An explanatory correlational design was used and the participants were assigned to different groups depending on their sex, age and education; once they met the inclusion criteria.

Data analysis: The sample was divided into two sub-samples in a random way. The first one was formed by 80% of the data and the second one by the remaining 20%. Sub-sample 1 was used to generate normative data, while sub-sample 2 was used to compare the model generated in sub-sample 1. The effects of demographic variables were evaluated by multiple regression analysis and included age, age<sup>2</sup>, schooling, education<sup>2</sup> and sex as predictor variables, as well as interactions with these variables. The scores of sub-sample 2 were transformed into  $z_i$  scores and percentiles from the normative data published in 2015 (Rivera and et al., 2015; Arango-Lasprilla and et al., 2015a and 2015b); that is, each direct score of sub-sample 2 was transformed to  $z_i$  scores and a percentile using two models: the first model resulting from this study and the second model published in 2015.

Results: Normative data for the three neuropsychological tests were obtained and organized in percentile tables. Linear effects of age and schooling were observed on the performance of the participants. No effects of sex on performance were found. Additionally, multiple regression models were obtained for each test and sub-test. When comparing the 2015 and 2020 models, it was found that the 2020 model explains a higher percentage of variance in all the tests and sub-tests compared.

Conclusion: Normative data without bias for age, education and sex were obtained in the three applied neuropsychological tests. In all the tests and sub-tests the effect of age and schooling was demonstrated. The use of multiple regression techniques produces normative data without bias for age, education and sex.

Keywords: Normative data, neuropsychological tests, multiple regression, adults.

## **Capítulo 1. Marco teórico**

### **1.1 Funciones ejecutivas.**

Las pruebas que se aplicaron en este trabajo evalúan distintos procesos cognoscitivos. Por ejemplo, la Prueba de Colores y Palabras de Stroop evalúa la flexibilidad y velocidad cognoscitiva, además de la atención selectiva, memoria de trabajo, entre otras funciones ejecutivas.

En el caso de la Prueba Modificada de Clasificación de Tarjetas de Wisconsin lo que evalúa es el razonamiento abstracto, la planeación estratégica, la respuesta a la retroalimentación, la flexibilidad cognoscitiva y el control de impulsos entre otras funciones ejecutivas.

La prueba de Dígito-Símbolo evalúa funciones como la atención dividida, el rastreo visual, la búsqueda visual, la velocidad visomotora y la memoria de trabajo, entre otras funciones ejecutivas.

#### **Definición**

La Función Ejecutiva (F. E.) es un conjunto de procesos cognoscitivos que involucran la selección, programación y regulación de la sensación y la motricidad. Incluye a los procesos de atención sostenida, la fluidez y la flexibilidad para generar nuevas respuestas, y la planeación y ejecución de actividades dirigidas a lograr un objetivo (Padilla, Téllez, Galarza, Téllez, Garza y Garza, 2016).

La mayoría de las definiciones de F. E. enumeran un grupo de procesos cognoscitivos que se denominan ejecutivos, como la flexibilidad cognoscitiva, la elección de objetivos, la planificación, el monitoreo, el uso de la retroalimentación, la solución de problemas, la

formulación de conceptos abstractos, el autocontrol, la autoconciencia, la inhibición, la atención, la memoria de trabajo, y la formulación de metas.

El termino F. E. ha tenido muchas conceptualizaciones, tantas y tan variadas que han llegado a referirse a un conjunto de procesos que han terminado por difuminarse y perderse teóricamente, restando operatividad y hasta validez al concepto. A continuación, se presentan algunas definiciones y al final se muestra una tabla en la que aparecen los componentes de la función ejecutiva propuestos por diferentes autores.

Según Soprano (2003) las F. E. son las habilidades para organizar y planificar una tarea, seleccionar apropiadamente objetivos, iniciar un plan y mantenerlo en la mente, inhibir distracciones, cambiar de estrategias, autorregular y controlar el curso de la acción y asegurar las metas.

Las F. E. son un grupo de procesos que aseguran la realización de un plan, son los mecanismos regulatorios de la actividad bajo el formato de un esquema de precisión y control; retoman los propósitos y demandan la información recolectada por la percepción, la memoria y la atención. De este modo, la actividad intelectual genera la línea de acción, la dirección de la atención mantiene el curso de la actividad, la memoria de trabajo permite tener presentes de manera simultánea varios procesos, la percepción aporta información precisa y rescata los parámetros relevantes de los datos, el lenguaje y el pensamiento simplifican el manejo de información compleja, y las FE son procesos integradores y reguladores que aseguran una ejecución ordenada y con sistemas de retroalimentación y control. Esta dinámica no es lineal, sino redundante, reverberante en ciclos de cadena corta y de cadena larga.

Para Valdez, Nava, Tirado, Frías y Corral (2005) las FE implican la programación y el autocontrol del comportamiento. Permiten la toma de decisiones, el autocontrol y la

solución eficiente de problemas. Las FE implican cuatro componentes: la iniciativa, la planeación, la acción y la verificación/corrección.

Por su parte Pineda (2000) considera las F.E. como un conjunto de habilidades cognoscitivas que permiten la anticipación y el establecimiento de metas, el diseño de planes y programas, el inicio de las actividades y operaciones mentales, la autoregulación y monitoreo de tareas, la selección precisa de los comportamientos, la flexibilidad del trabajo cognoscitivo y su organización espacio-temporal para resolver problemas.

Las FE son los procesos mentales mediante los que resolvemos deliberadamente problemas internos y externos. La meta de las FE es solucionar estos problemas de una forma eficaz y aceptable para la persona y la sociedad (Papazian, Alfonso y Luzondo, 2006).

Estévez-González, García-Sánchez y Barraquer-Bordas (2000) le dan un sentido más temporal a las FE dado que las conceptualizan como organizadores cronológicos de la conducta. De esta manera se asocian ideas simples y se combinan hacia la resolución de problemas de alta complejidad.

Lezak (1995) define las FE como las capacidades cognitivas esenciales para llevar a cabo una conducta eficaz, creativa y aceptada socialmente.

Sohlberg y Mateer (2001) consideran que las FE abarcan una serie de procesos cognoscitivos entre los que destacan la identificación de objetivos, planeación y organización de conductas para lograr objetivos y el monitoreo de actividades para conseguir metas. En la Tabla 1 se presentan algunos procesos cognoscitivos implicados en las FE.

Tabla 1. Algunos procesos implicados en las funciones ejecutivas.

PROCESO	AUTOR
Organizar y planificar una tarea	(Estévez-González, García-Sánchez y Barraquer-Bordas, 2000) (Padilla, et al, 2016) (Pineda, 2000) (Sohlberg y Mateer, 2001) (Soprano, 2003)
Seleccionar objetivos	(Estévez-González, García-Sánchez y Barraquer-Bordas, 2000) (Padilla, et al, 2016) (Pineda, 2000) (Sohlberg y Mateer, 2001) (Soprano, 2003)
Iniciar un plan y mantenerlo en la mente	(Pineda, 2000) (Valdez, et al., 2005) (Soprano, 2003)
Inhibir distracciones	(Soprano, 2003).
Cambio de estrategias	(Padilla, et al, 2016) (Pineda, 2000) (Valdez, et al., 2005) (Soprano, 2003)
Autorregular y controlar el curso de la acción	(Padilla, et al, 2016) (Pineda, 2000) (Valdez, et al., 2005) (Sohlberg y Mateer, 2001) (Soprano, 2003)
Asegurar metas	(Padilla, et al, 2016) (Soprano, 2003)
Anticipación	(Pineda, 2000) (Sohlberg y Mateer, 2001)
Organización espacio-temporal	(Estévez-González, García-Sánchez y Barraquer-Bordas, 2000) (Pineda, 2000)
Solución de problemas	(Papazian, Alfonso y Luzondo, 2006) (Valdez., y et al., 2005)
Retroalimentación.	(Sohlberg y Mateer, 2001) (Valdez y et al., 2005)

*Elaboración propia.*

Miyake et al., (2000) consideran que hay 3 funciones básicas para la correcta ejecución de tareas: a) el cambio entre tareas (shifting), b) actualización (updating) y c) monitoreo de la memoria de trabajo y de la inhibición de respuestas dominantes. Para estos autores las dos últimas funciones son la base de la comunalidad de las FE.

A manera de conclusión pudiéramos decir que las F.E. son un conjunto de procesos que permiten la autorregulación de la conducta a través de la elaboración de estrategias de comportamiento en las que se involucran la elaboración de planes de conducta, su ejecución y verificación hasta la consecución de objetivos. En las funciones ejecutivas están involucrados procesos como la memoria de trabajo (verbal y no verbal), la inhibición, la anticipación de resultados, el cambio en el foco de atención, el uso de la retroalimentación y

el seguimiento de reglas sociales. De esta manera la persona es capaz de adaptarse a los requerimientos que le demande el ambiente social y podrá solucionar problemas de acuerdo a un plan que debe ser verificado y corregido si es el caso.

## **1.2 Pruebas neuropsicológicas que miden las funciones ejecutivas**

### **1.2.1 La Prueba de Colores y Palabras (Stroop).**

La Prueba de Colores y Palabras, también llamada “Stroop Test”, fue una prueba creada en 1935 por John Ridley Stroop (McLeod, 1991; Stroop, 1935). La Prueba de Stroop mide la facilidad con la cual una persona es capaz de mantener un objetivo en mente y suprimir una respuesta habitual por otra respuesta menos familiar (Strauss, Sherman y Spreen, 2006). También es una prueba que evalúa la flexibilidad y la velocidad cognoscitiva, el control atencional (Banich et al., 2000), atención selectiva (Ben-David et al., 2011; Zhang et al., 2013) y las F.E. (Heflin et al., 2011).

Existen diferentes versiones de la prueba, la versión original (Stroop, 1935) estaba compuesta por tres tarjetas blancas, cada una con 10 líneas de cinco ítems cada una. En la primera parte la persona debía de leer los nombres de colores (azul, rojo, verde, café y púrpura) impresos en negro. En la segunda parte debía leer el nombre de los colores (los mismos de la primera parte) impresos con una tinta distinta al nombre del color escrito. En la tercera parte debía decir el color de unos cuadrados. En la cuarta parte se le daban las tarjetas utilizadas en la segunda parte y debía de decir el color en el que están escritas las palabras, ignorando la palabra que está escrita.

Otra de las versiones de la prueba es el Victoria Stroop Test (Regard, 1981) que consta de solo 24 ítems en cada una de las tres tarjetas que lo componen. En la primera tarjeta la persona debe nombrar colores; en la segunda, leer palabras neutras (nombres de colores); y en la tercera decir el color de palabras impresas en colores contrastantes.

La versión utilizada en este trabajo es la de Golden (2007) que consta de tres páginas en las que hay 100 ítems organizados en 5 columnas de 20 ítems cada una. En la primera se presentan tres nombres de colores (azul, verde y rojo) escritas en tinta negra (Figura 1). En la segunda página se presentan grupos de cuatro letras equis (XXXX) impresas al azar en cada uno de los tres colores (azul, verde o rojo) y la persona debe nombrar el color de cada grupo de letras X (Figura 2). En la tercera página se presentan las mismas palabras que se presentaron en la primera, pero impresas en los mismos colores (mismo orden) de los grupos de letras X que se presentaron en la segunda página y la persona debe de decir el color en el que están escritas estas palabras sin prestar atención (inhibiendo) a la palabra escrita, sino al color en el que está impresa (Figura 3). En cada parte de la prueba se les da un tiempo límite de 45 segundos y se registra el número de ítems correctos que logra en cada una.

CAFÉ	ROSA	LILA	CAFÉ	ROSA
LILA	LILA	CAFÉ	ROSA	LILA
ROSA	CAFÉ	ROSA	LILA	CAFÉ
LILA	ROSA	CAFÉ	CAFÉ	ROSA

Figura 1. Ejemplos de las primeras cuatro líneas de estímulos de la primera página de la prueba de Colores y Palabras de Stroop. Los estímulos han sido cambiados de acuerdo con la guía de International Test Commission (2014)

XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX
XXXX	XXXX	XXXX	XXXX	XXXX

Figura 2. Ejemplos de las primeras cuatro líneas de estímulos de la segunda página de la prueba de Colores y Palabras de Stroop. Los estímulos han sido cambiados de acuerdo con la Guía de International Test Commission (2014)

CAFÉ	ROSA	LILA	CAFÉ	ROSA
LILA	LILA	CAFÉ	ROSA	LILA
ROSA	CAFÉ	ROSA	LILA	CAFÉ
LILA	ROSA	CAFÉ	CAFÉ	ROSA

Figura 3. Ejemplos de las primeras cuatro líneas de estímulos de la tercera página de la prueba de Colores y Palabras de Stroop. Los estímulos han sido cambiados de acuerdo con la Guía de International Test Commission (2014)

Algunos estudios han reportado que existe una relación entre la ejecución de esta prueba y la activación de áreas específicas del cerebro. Por ejemplo, se ha demostrado con estudios de Imagen por Resonancia Magnética (MRI, por sus siglas en inglés), la relación entre el rendimiento en el Stroop y las áreas parietales y occipitales (Herd, Banich, y O'Reilly, 2006), parietal inferior y frontal dorsolateral (Banich, et al. 2000), área motora suplementaria, giro pre-central, ínsula anterior (Sachs, Kaplan, Der Sarkissian, y Habibi, 2017), y el estriado dorsal y tálamo (Carley et al., 2015).



También se ha demostrado a través de la MRI funcional que existe una relación entre (fMRI) entre la activación de diferentes zonas del lóbulo frontal y la ejecución en la prueba de Stroop, lo que apoya la idea de que el Stroop es una prueba que evalúa las F.E. (Heflin et al, 2011).

Al ser el Stroop una de las pruebas más ampliamente utilizadas en el mundo para evaluar las F. E., es común encontrar que esta haya sido utilizada en la evaluación neuropsicológica de los problemas atencionales y déficits en las F. E. de personas con una amplia variedad de alteraciones neurológicas y psiquiátricas, entre las que se encuentran: la enfermedad de Alzheimer (Duchek et al., 2013; Melrose, et al., 2017), esclerosis múltiple (Vitkovitch, Bishop, Dancey, y Richards, 2002), Parkinson (Bezdicek et al., 2015), traumatismos craneoencefálicos (Ben-David, Nguyen, y van Lieshout, 2011; Guise, Thompson, Greve, Bianchini, y West, 2014), esquizofrenia (Perlstein, Carter, Barch, y Baird, 1998) depresión, (Hammar, et al., 2010), accidentes cerebrovasculares, (Negahban, Ebrahimzade, y Mehravar, 2017; Park, y Lee, 2019).

Igualmente, la Prueba de Stroop ha sido ampliamente utilizada en población infantil y adolescente (Rivera et al., 2017; Beltrán y Solís, 2012;), adultos (Amato et al., 2006; Peña-Casanova et al., 2009; Rivera et al., 2015) y en población de la tercera edad (Amato et al., 2006; Anstey et al., 2000; Peña-Casanova et al., 2009; Vogel et al., 2013).

Con respecto a las propiedades psicométricas, se han realizado estudios para revisar la confiabilidad de la prueba, por ejemplo, en uno de los primeros estudios sobre la confiabilidad de diferentes versiones de la prueba, Siegrist (1997) encontró índices de .84 a .91 ( $p < 0.001$ ) en las versiones que utilizó. Por otro lado, Van der Elst, Van Boxtel, Van Breukelen y Jolles (2008) utilizando la versión de Hammes encontraron que la confiabilidad

test-retest era aceptable (0.68,  $p < 0.001$ ). En otro estudio realizado en Colombia en 2016; Rodríguez, Pulido y Pineda encontraron índices de confiabilidad de .78 a .91 ( $p < 0.001$ ).

#### *1.2.1.1 Variables que influyen en la ejecución.*

Diferentes estudios se han realizado con el objetivo de determinar cuáles son las variables que están relacionadas con el desempeño de la prueba. Por ejemplo, se ha encontrado que la escolaridad es una variable muy importante para el desempeño de esta prueba, puesto que a mayor escolaridad el rendimiento en la prueba suele ser mucho mejor (Beltrán y Solís, 2012; Peña-Casanova et al., 2009; Rivera et al., 2015; Vogel et al., 2013;).

La edad es otra variable que también tiene efectos directos sobre la ejecución en la Prueba de Stroop (Peña-Casanova et al., 2009; Strauss et al., 2006). Es común encontrar que las personas más jóvenes presenten un mejor rendimiento en la prueba en comparación con el rendimiento de personas mayores (Armengol, 2002; Oliveira, Mograbi, Gabrig, y Charchat-Fichman, 2016; Peña-Casanova et al., 2009; Vogel et al., 2013). En un estudio realizado por Vogel et al., (2013) en el que se aplicó una versión danesa del Stroop a un grupo de adultos mayores saludables se encontró que la edad tuvo un efecto favorable en los puntajes Color-Palabra, siendo los más jóvenes los que presentaron una ejecución más rápida que los más viejos, lo mismo ocurrió con las correlaciones entre Color-Palabra con edad y con escolaridad.

El género también se ha reportado como un factor que puede determinar la ejecución en el Stroop, aunque en ocasiones estas diferencias entre género no son significativas, se ha reportado que en general las mujeres obtienen mejores puntajes en la modalidad palabras (Moering et al., 2004; Strickland et al., 1997; Van der Elst et al., 2006).

Dado que la Prueba de Stroop depende de que la persona sepa leer (hay que recordar que en la segunda y tercera parte no se le pide a la persona que lea, sino que denomine

colores), se han realizado estudios para evaluar qué tan determinante es esta habilidad de lectura en la ejecución de la prueba (Beltrán y Solís, 2012; Kapoula et al. 2010; Proulx y Elmasry, 2015; Savitz y Jansen, 2003). Beltrán y Solís (2012) evaluaron niños y adolescentes y reportaron que la escolaridad (y muy probablemente la lectura) tienen un efecto significativo sobre la ejecución de la prueba. El caso contrario ocurre cuando el test se le aplica a personas con dificultades de lectura (Savitz y Jansen, 2003).

#### *1.2.1.2 Estudios con datos normativos.*

En la actualidad existen datos normativos para la prueba de Stroop en diferentes países tales como Francia, (Roy, Kefi, Bellaj, Fournet, Le Gall, y Roulin, 2018), España, (Llinas-Regla, Vilalta-Franch, López-Pousa, Calvo-Perxas, y Garre-Olmo, 2013), Italia (Brugnolo et al., 2016), Brasil (Oliveira et al., 2016; Zimmerman, Cardoso, Trentini, Grassi-Oliveira, y Fonseca, 2015), Colombia (Beltrán y Solís, 2012; Rodríguez, Pulido, y Pineda, 2016), China (Luo, Proctor y Weng, 2015), México (Armengol, 2002), Portugal (Martins et al., 2005) y Estado Unidos de Norteamérica (Moran y Yates, 2011). Cabe aclarar que el único estudio que utilizó análisis de regresión multivariado fue el de Brugnolo: sin embargo, no usó los residuos estandarizados para la generación de los datos normativos.

Con respecto a población de habla hispana, la prueba ha sido baremado en 347 personas con edades comprendidas entre los 50 a 90 años en España (Peña-Casanova et al., 2009). También se poseen datos normativos de 5402 adultos (Rivera et al., 2015; Olivera et al., 2015), de edades comprendidas entre los 18 y 90 años en 12 países latinoamericanos (Colombia, Chile, Perú, México, Honduras, Paraguay, Argentina, Salvador, Guatemala, Bolivia, Cuba y Puerto Rico).

### 1.2.2 Prueba Modificada de Clasificación de Tarjetas de Wisconsin.

La Prueba de Clasificación de Tarjetas de Wisconsin fue desarrollada por Grant y Berg en 1948, fue diseñada originalmente para evaluar el razonamiento abstracto y la habilidad para cambiar de estrategia cognoscitiva en respuesta a cambios en las contingencias ambientales y también se utilizó para evaluar la solución de problemas, la planeación estratégica, el uso de la retroalimentación ambiental para cambiar y la inhibición de respuestas (Damakis, 2003).

La prueba original consta de 128 tarjetas en las que se presentan de 1 a 4 figuras (triángulos, estrellas, cruces o círculos) en 4 posibles colores (rojo, verde, amarillo o azul). Además, hay 4 tarjetas clave para que la persona pueda clasificar cada una de las tarjetas, de manera individual, de acuerdo con alguna característica que comparta con cualquiera de las tarjetas clave ya sea color, forma o número (Berg, 1948; Grant y Berg, 1948). La persona evaluada debe de colocar cada tarjeta debajo de alguna de las tarjetas clave y el evaluador debe retroalimentar la elección de la persona diciendo "CORRECTO" o " INCORRECTO". La secuencia de categorización es color, luego forma y finalmente número; una vez que se logra la tercera categoría se repiten las categorías en el mismo orden. Después de diez respuestas correctas consecutivas el criterio de clasificación cambia sin avisarle al evaluado y se debe de ajustar la decisión. La prueba termina cuando la persona logra seis categorías correctas o cuando se terminan las 128 tarjetas. En realidad, son dos mazos iguales con las mismas 64 tarjetas, que se presentan en el mismo orden, es decir la tarjeta uno y la 65 son iguales, lo mismo pasa con la dos y la 66 y así con todas las demás. Asimismo, existe la versión corta de esta prueba en la que se mantienen las mismas 4 tarjetas clave pero solo se presentan 64 tarjetas. El procedimiento de aplicación y de calificación es el mismo.

La versión Modificada de la Prueba de Clasificación de Tarjetas de Wisconsin (Nelson, 1976) consta de las mismas 4 tarjetas clave, pero solamente hay 48 tarjetas para clasificar. A esta versión se retiraron 80 tarjetas de la versión original que comparten simultáneamente más de un atributo de clasificación con cualquiera de las 4 tarjetas clave. Esto permite que el evaluador no tenga dudas sobre cuál es el criterio que está utilizando el evaluado al colocar una tarjeta que comparte varias características clasificatorias con alguna de las tarjetas clave. Además de esta diferencia sustancial, existen otras en la aplicación de la prueba: al paciente se le dice que el criterio de clasificación de las tarjetas puede cambiar y que cuando eso ocurra se le avisará; lo que reduce la frustración y por cierto, reduce los errores perseverativos típicos de los cambios de categorización. Además, el criterio para cambiar el principio de clasificación es lograr 6 respuestas correctas consecutivas. Otra diferencia es que las categorías a clasificar no están predeterminadas, sino que la persona evaluada decreta el orden de las primeras tres categorías y cuando la tercera categoría es lograda, el orden de las primeras tres categorías se repite en las últimas tres. En la Figura 4 se muestra la manera en la que se le presentan las tarjetas clave a la persona evaluada.

Los elementos a registrar son la cantidad de categorías logradas correctamente, el número de errores totales, el número de errores perseverativos y el porcentaje de errores perseverativos. El Wisconsin es una prueba que se ha utilizado con una serie de poblaciones pediátricas tales como pacientes con discapacidad intelectual (Rockers, Ousley, Sutton, Schoenberg, Coleman, Walker y Cubells, 2009), autismo (Landry y Al-Taie, 2016; Ozonoff, 1995), ADHD (Ter-Stepanien, Grizenko, Cornish, Talwar, Mbekou, Schmitz y Joobar, 2017; Scheres et al., 2004), niños con síndrome fetal de alcohol (Vaurio, Riley, y Mattson, 2008), niños bilingües, (Vega y Fernández, 2011; Villamizar y Guevara, 2013), y niños con inteligencia media y superior (Arffa, Lovell, Podell, y Goldberg, 1998).

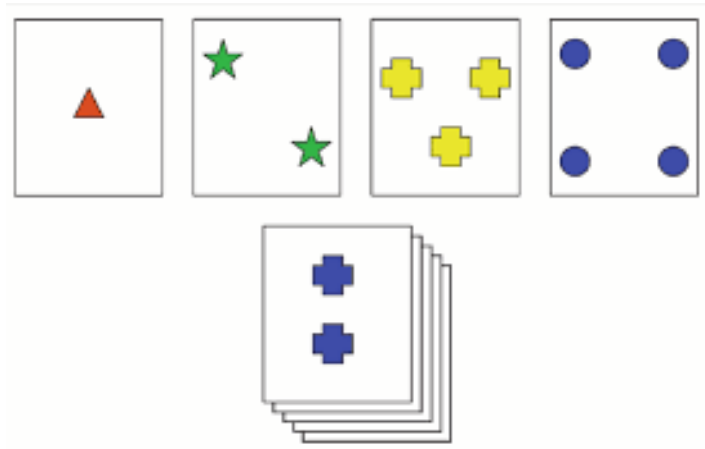


Figura 4. Se muestra el arreglo de las tarjetas clave o guía de la Prueba Modificada de Clasificación de Tarjetas de Wisconsin.

La prueba también es una prueba ampliamente utilizada en personas adultas con esquizofrenia (Rybakowski, Borkowska, Czerski, Kapelski, Dimitrza-Weglarz, y Hauser, 2005), diabetes (Christman, Vahorsdall, Pearlson, Hill-Briggs, y Schretlen, 2010; Fernandes-Lopes y De Lima-Argimon, 2010), trauma de cráneo (Sherer, Nick, Millis y Novack, 2003; Stuss et al., 2000), esclerosis (Kuccu, Kandemir, Unal, Topcular y Kirbas, 2012), tumores (Soto, Cairo, Marrero y Gonzalez, 2016), Enfermedad de Alzheimer (Takeda et al., 2010), accidentes cerebro vasculares (Greve, Bianchini, Hartley, y Adams, 1999; Jodzio, y Biechowska, 2010) entre otros.

#### *1.2.2.1 Variables que influyen en la ejecución de la prueba.*

Diferentes estudios han encontrado que la edad, la escolaridad, el género y el coeficiente intelectual son variables que pueden influir en el desempeño de la prueba (Arango-Lasprilla et al., 2015; Cianchetti, Corona, Foscoliano, Contu y Sannio-Fancell, 2007; Fristoe, Salthouse y Woodward, 1997; Lin, Chan, Zheng, Yang, y Wang, 2007)

Con respecto a la edad, Cianchetti, Corona, Foscoliano, Contu y Sannio-Fancell, (2007) realizaron un estudio con un grupo de niños de 4 a 13 años en los que encontraron que el rendimiento en cada una de las puntuaciones del Wisconsin fue mucho mejor en los niños mayores. En otro estudio Fristoe, Salthouse y Woodward (1997) compararon la ejecución de adultos jóvenes (18 a 38 años de edad) y adultos mayores (60 a 86 años) con respecto a la cantidad de errores sencillos y errores de perseveración, en 97 adultos. Los autores encontraron que los adultos mayores tenían una ejecución en la prueba de Wisconsin con un mayor número de errores totales y de errores perseverativos. Así mismo, Lin, Chan, Zheng, Yang y Wang (2007) aplicaron en China el WCST a dos grupos de adultos (60 a 70 años de edad y de más de 70 años de edad) y encontraron diferencias entre ambos grupos en la cantidad de errores perseverativos y en la cantidad de categorías logradas con una mejor ejecución en el grupo de menor edad.

Finalmente, en un estudio realizado en Latinoamérica en una muestra de 5,402 adultos evaluados en 12 países, también se reportaron efectos de la edad (Arango-Lasprilla et al., 2015), de forma que, en la mayoría de los países las puntuaciones en la prueba empeoran conforme incrementa la edad.

Con respecto a la escolaridad se ha encontrado que el nivel de educativo tiene efectos sobre la ejecución del M-WCST. Arango-Lasprilla et al., (2015) encontraron que las puntuaciones en la prueba incrementan de forma lineal en función de los años de escolaridad. Es decir, a mayor nivel educativo mejor rendimiento en la prueba.

Con respecto a la relación entre el género y desempeño en la prueba no se han encontrado resultados concluyentes. Al parecer la gran mayoría de estudios coinciden con afirmar que no existe una relación directa entre ambos (Arango-Lasprilla et al., 2015; Caffarra et al., 2010; Obonsawin, 1999).

Respecto a la influencia de la inteligencia sobre los resultados de la prueba, son pocos los estudios que la han evaluado. En un estudio realizado por Arffa, Lovell, Podell y Goldberg (1998) se aplicaron el WCST y el WISC a una muestra de 50 niños. La mitad de la muestra tenía un coeficiente intelectual (CI) de entre 110 y 129, mientras que la otra mitad tenían un CI mayor de 130. Su resultado principal fue que el grupo con el CI más alto obtuvo mejores puntuaciones en cada uno de los índices registrados de la prueba. En otro estudio realizado por Ardila, Pineda y Rosselli (2000) aplicaron el WCST y el WISC-R a 50 niños y encontraron una correlación negativa entre el CI Verbal, el CI Total y los errores perseverativos.

#### *1.2.2.2 Estudios con datos normativos.*

El M-WCST es un test que ha sido ampliamente estudiado y baremado en diferentes países (Arango et al., 2015c y 2015d; Caffarra, Vezzadini, Dieci, Zonato y Vanneri, 2004; Cianchetti, Corona, Foscoliano, Contu y Sannio-Fancello, 2007; Obonsawin, et al., 1999; Wang et al. 2011), poblaciones con desorden por déficit de atención (Romine et al., 2004); desórdenes de conducta (Morgan y Lilienfeld, 2000); autismo (Ozonoff, 1995), epilepsia de lóbulo temporal (Igarashi et al., 2002), por grupos de edad como por ejemplo a niños italianos (Cianchetti, Corona, Foscoliano, Contu, y Sannio-Fancello, 2007), en niños y adolescentes (Diforio, Walker y Kestler, 2000), adultos jóvenes con el síndrome de delección 22q11.2 (Rockers et al., 2009) y adultos (Arango-Lasprilla y et al., 2015b; 2015c; Caffarra, Vezzadini, Dieci, Zonato y Vanneri, 2004; Obonsawin, et al., 1999). A continuación, se describen los principales estudios de baremación que se han realizado a nivel internacional.

En un estudio Obonsawin, Crawford, Page, Chalmers, Low y Marsh (1999) aplicaron el M-WCST a 146 personas de entre 16 y 75 años de edad con el propósito de obtener datos normativos para el Reino Unido. El estudio generó tablas de percentiles para el número de



categorías logradas, el total de errores, errores perseverativos y porcentaje de errores perseverativos.

Por otra parte, Caffarra, Vezzadini, Dieci, Zonato y Vanneri, en 2004 aplicaron en Italia el M-WCST a 248 adultos de entre 20 a 90 años de edad con el propósito de presentar generar datos normativos. En este estudio los investigadores encontraron un efecto significativo de la edad y la escolaridad en las puntuaciones del número total de categorías completadas y errores perseverativos.

También en Italia, Cianchetti, Corona, Foscoliano, Contu y Sannio-Fancello (2007) con el objetivo de generar datos normativos aplicaron el M-WCST a 1,126 niños de cuatro a 13 años de edad. Los resultados fueron presentados en forma de tablas con promedios y desviaciones estándar para cada grupo de edad (4-13 años). Los investigadores no encontraron diferencias significativas entre niños y niñas, pero si encontraron que a medida que la edad incrementa, estos mejoran en algunos de los parámetros de la prueba, aunque esta mejoría fue menos relevante en las edades más altas, especialmente en el número de categorías completadas.

Arango-Lasprilla y et al., (2015b; 2015c) aplicaron el M-WCST a un grupo de 5,402 personas entre los 18 y los 90 años de edad en 12 países de Latinoamérica. Los datos se presentaron por país de acuerdo a aquellas variables que resultaron estar significativamente asociadas al rendimiento de las pruebas. En general se encontró que edad y la escolaridad estaban asociadas al rendimiento en la prueba. Mientras que no encontraron que la variable género tuviera un efecto significativo en la ejecución de la prueba.

### 1.2.3 Prueba de modalidades símbolo-dígito (SDMT).

La Prueba de Símbolo-Dígito (SDMT) fue diseñada por Smith en 1968 inspirado en una parte del Test Army Beta. La SDMT se diseñó para evaluar el seguimiento y la

exploración visual y la velocidad de procesamiento, pero también sirve para evaluar la atención dividida, la velocidad motora y la memoria de trabajo (Lezak, Howieson, y Loring, 2004; Shum, McFarland, y Bain, 1990; Strauss, Sherman y Spreen, 2006).

La Prueba consta de una clave en la que aparecen en una línea doble de cuadrados, una serie de nueve símbolos abstractos a los que les corresponden un número del uno al nueve. El símbolo aparece en la parte superior y el número en la inferior (Smith, 2002). Debajo de esta clave aparece de nuevo la misma línea doble con cuadrados, pero en esta solo aparecen los símbolos; y la persona debe de escribir el número que le corresponde a cada uno. Los primeros 10 espacios a llenar son de ensayo para que la persona practique la tarea y los subsecuentes son los que se calificarán. A la persona se le pide que inicie cuando el evaluador lo indique y cuando transcurren 90 segundos se le pide que se detenga. El puntaje que se toma en cuenta es la cantidad de respuestas correctas que la persona hace, y el máximo que se puede obtener es 110 (Smith, 2002; Smith, 2013). En la Figura 5 se muestran la clave y los ejemplos del SDMT.

ID: \_\_\_\_\_

Puntaje: \_\_\_\_\_

**CLAVE**

(	÷	┌	┐	└	>	+	)	÷
1	2	3	4	5	6	7	8	9

**EJEMPLOS**

(	└	÷	(	┌	>	÷	┐	(	>	÷	(	>	┐	÷

┐	>	(	÷	└	>	┌	(	÷	>	÷	┐	┌	)

┐	└	+	)	(	┌	+	┐	)	└	÷	÷	┌	┐	+

÷	┐	└	(	>	┐	(	└	>	+	÷	)	┌	>	┐

Figura 5. Ejemplo de la clave utilizada en SDMT. Se muestran los diez ejemplos que sirven de práctica a la persona evaluada y la mitad de las líneas en las que tiene que trabajar.

El SDMT se puede aplicar de manera escrita o en forma oral. Según el manual, su aplicación es para personas desde los 8 hasta los 78 años de edad. La aplicación puede ser de manera individual o de forma colectiva, existen versiones computarizadas de la prueba. Así mismo existen otras variantes, ya sea escrita o computarizada, en las que los símbolos son distintos. Incluso hay una versión modificada utilizada por Gonzalez, Whitfield, West, Williams, Lichtemberg y Jackson (2007) que está compuesta por dos secciones de 50 estímulos cada una y el tiempo que se le otorga al paciente para completar la tarea es de 45 segundos en cada parte.

La prueba de SDMT ha sido aplicada a diferentes poblaciones, tales como personas con esclerosis (Drake, et al., 2010; Hubbard, y Motl, 2015; Schependom, et al., 2015; Settle,

Robinson, Kane, Maloni y Wallin, 2015;), trauma de cráneo (Dymoswski, Owens, Pansford y Willmott, 2015), infartos cerebrales (Kho, et al., 2011), HIV (Benedict, Mezhir, Walsh y Hewitt en 2000), diabetes (Chaytor, Barbosa-Leiker, Ryan, Germine, Hirsh y Weinstock, 2019) y enfermedad de Parkinson (Lee, Falkmer, y Chee, 2011; Sarno et al., 2019).

#### *1.2.3.1 Variables que influyen en la ejecución de la prueba.*

Se ha encontrado que la edad es un factor que influye en el rendimiento del SDMT. Esencialmente lo que se ha reportado es que durante la niñez los puntajes obtenidos van incrementando hasta la edad adulta y luego al ir envejeciendo los puntajes en SDMT empiezan a disminuir (Arango-Lasprilla et al., 2015a y 2017; Burggraaff, Knol y Uitdehaag, 2017; Cancela, Ayán y Varela, 2012; Peña-Casanova, et al., en 2009; Vanotti, et al., 2015).

En relación a la influencia del género en el desempeño de la prueba, no se ha encontrado que exista una asociación entre ambos (Arango-Lasprilla et al., 2015b y 2017b; Burggraaff, Knol y Uitdehaag, 2017; Cancela, Ayán y Varela, 2012; Peña-Casanova, et al., en 2009; Vanotti, et al., 2015). Sin embargo, Kiely, Butterworth, Watson y Woodens (2014) aplicaron en Australia el SDMT a 14,456 personas y reportan que las mujeres tuvieron una mejor ejecución en la prueba. A partir de la revisión hecha por ellos concluyen que esta diferencia podría ser debida a que las mujeres tienen mejor desempeño que los hombres, en la codificación verbal para símbolos abstractos. Sin embargo, en los demás estudios reportados por otros investigadores estas diferencias no han sido encontradas.

Los años de escolaridad completados han demostrado ser un factor determinante en los puntajes obtenidos en el SDMT. Independientemente del género o del país en el que viva la persona evaluada; la educación ha demostrado tener un impacto favorable en la ejecución de las personas. En el caso del SDMT se ha reportado que a mayor cantidad de años de estudio completados, es mejor el puntaje obtenido en la prueba. Esto se ha observado incluso

en los grupos de adultos mayores en los que la escolaridad mejora la ejecución en esta prueba (Arango-Lasprilla et al., 2015b y 2017b; Burggraaff, Knol y Uitdehaag, 2017; Cancela, Ayán y Varela, 2012; Kiely, Butterworth, Watson y Woodens, 2014; Peña-Casanova, et al., en 2009; Vanotti, et al., 2015).

En general los estudios de datos normativos que se han realizado no reportan efectos de la cultura o de la pertenencia a algún grupo étnico sobre la ejecución en el SDMT (Arango-Lasprilla y et al., 2015b y 2017b; Burggraaff, Knol y Uitdehaag, 2017; Cancela, Ayán y Varela, 2012; Kiely, Butterworth, Watson y Woodens, 2014; Peña-Casanova, et al., 2009; Vanotti, et al., 2015). Sin embargo, otros estudios si han encontrado que ciertos grupos étnicos tales como los afrocaribeños presentan peores puntuaciones en el SMDT en comparación con afroamericanos y blancos no latinos (Gonzalez et al., 2007). Un estudio realizado en Australia por Kiely, Butterworth, Watson y Woodens, (2014) en una muestra que incluyó personas que no habían nacido en Australia y cuyo idioma natal no era el inglés, personas aborígenes y personas no aborígenes que habían nacido en Australia, reportaron que los participantes aborígenes tenían una puntuación por debajo de la obtenida por los nacidos fuera de Australia, y estos últimos obtenían un puntaje más bajo que los nacidos en Australia.

#### 1.2.3.1 Estudios con datos normativos.

El SMDT es una de las pruebas neuropsicológicas que se ha estudiado más ampliamente a nivel internacional y en la actualidad se dispone de datos normativos de esta prueba en poblaciones de diferentes países del mundo. A continuación, se detallan algunos de los principales estudios que se han llevado a cabo con esta prueba.

Peña-Casanova, et al., (2009) aplicaron el SDMT a una muestra de 354 españoles de entre 50 y 90 años de edad, con escolaridad elemental hasta la superior, los autores encontraron que el género no tuvo efecto sobre la ejecución en la prueba; pero en el caso de

la edad, esta correlacionaba negativamente con los puntajes y que la escolaridad se relacionaba positivamente con la ejecución.

Por su parte, Cancela, Ayán y Varela (2012) aplicaron el SDMT a 321 pacientes de un hospital geriátrico en España. La muestra tenía participantes de entre 55 y 90 años de edad y escolaridades desde cero hasta 20 años completados. En esta muestra y dado que eran de un hospital geriátrico algunas personas tenían las siguientes condiciones: deterioro cognoscitivo leve (18.2 %), hipertensión (32.1%), diabetes (14.6%) y artrosis (13.1%). Los autores encontraron que la edad y la escolaridad estaban relacionadas con el desempeño en la prueba, sin embargo, no se encontró influencia del género.

En Argentina, Vanotti, et al., (2015) aplicaron una versión oral del SDMT a 297 hombres y mujeres con edades de entre 19 a 70 años y con escolaridades desde 3 a 21 años completados. No presentan tablas con datos estandarizados o percentiles, sin embargo, reportaron que el género no marcó diferencia en las puntuaciones obtenidas, pero hacen notar que a mayor edad se obtenían puntuaciones menores y que la alta escolaridad (en años completados) mejoraba la ejecución.

Kiely, Butterworth, Watson y Woodens (2014) administraron el SDMT a 14,456 australianos de ambos sexos, con edades de entre 15 y 100 años. Ellos encontraron que la edad, el género, la escolaridad, el bagaje cultural y el estado de salud estaban significativamente asociados con los puntajes de la prueba.

Arango-Lasprilla y et al., (2015) realizaron un estudio con el propósito de obtener datos normativos en 12 países de Latinoamérica ( Argentina, Bolivia, Chile, Colombia, Cuba, El Salvador, Guatemala, Honduras, México, Paraguay, Perú y Puerto Rico) aplicaron el SMDT a un grupo de 5402 adultos de 18 a 90 años con diferentes años de escolaridad. Estos investigadores generaron datos normativos por país y encontraron que la edad y la

escolaridad están asociadas a las puntuaciones obtenidas por las muestras de los diferentes países y que el género no influye en la ejecución.

Por su parte, Burggraaff, Knol y Uitdehaag en 2017, aplicaron, en Alemania, el SDMT a 96 personas sanas y a 157 pacientes con esclerosis múltiples y generaron datos normativos. No encontraron influencia del género, ni de la escolaridad sobre la ejecución en la prueba. Solamente encontraron en las personas sanas que la edad influía en los puntajes obtenidos.

## **Capítulo 2. Planteamiento del problema**

### **2.1. Planteamiento del problema**

La neuropsicología es un área de la psicología que se encarga de estudiar la relación entre el funcionamiento cerebral y el comportamiento humano, tanto en personas sanas como en personas con daño cerebral (Portellano, 2005). Una de las labores más importantes en el quehacer neuropsicológico es el proceso de evaluación y diagnóstico, el cual consiste en un proceso mediante el cual el neuropsicólogo realiza un estudio detallado del funcionamiento cognitivo, emocional y funcional de la persona (Goldstein y McNeil, 2004; Hebben y Milberg, 2011). Para ello una de las principales herramientas que el neuropsicólogo suele utilizar en su práctica clínica son las pruebas neuropsicológicas, los cuales deben de poseer la suficiente validez y confiabilidad para medir cada uno de los aspectos que se quieren evaluar. Igualmente, dichas pruebas deben de poseer datos normativos con los cuales se pueda comparar el desempeño de pacientes con aquellas personas de sus mismas características socioculturales y así saber si el rendimiento del paciente en una tarea específica se encuentra o no en el rango normal esperado (Hebben y Milberg, 2011; Lezak, Howieson y Loring, 2004; Strauss, Sherman, y Spreen, 2006)

En América Latina, la neuropsicología es una disciplina muy joven y sus inicios se remontan a la década de los años 70 (Arango-Lasprilla, et al., 2017). La gran mayoría de neuropsicólogos latinoamericanos durante muchos años han venido trabajando con pruebas neuropsicológicas que no han sido adaptadas o baremadas en población latinoamericana. Por tal motivo, durante muchos años el proceso de evaluación y diagnóstico se ha llevado a cabo a partir de la interpretación de las puntuaciones brutas o mediante la utilización de baremos



creados en otros países (Arango-Lasprilla, et al., 2015; Arango-Lasprilla, Stevens, Morlett-Paredes, Ardila, y Rivera, 2017).

Durante los últimos años se ha realizado un gran esfuerzo por contar cada vez más con pruebas neuropsicológicas que posean datos normativos en población latinoamericana (Beato, Amaral-Carvalho, Guimaraes, Tumas, Souza, Oliveira, y Caramelli, 2012; Butman, Allegri, Harris, y Drake, 2000; Magalhaes, y Hamdan, 2010; Ostrosky-Solís, Ardila, y Rosselli, 1999; Ostrosky-Solís, Dávila, Ortiz, Vega, Ramos, de Celis., et al. 1999; Rosselli, Ardila, Florez, y Castro, 1990). Sin embargo, todavía existen muchas pruebas de no poseen datos normativos baremados en esta población. En un estudio realizado por Arango-Lasprilla, et al., (2017) se encontró que algunas de las principales pruebas neuropsicológicas más comúnmente utilizadas en población latinoamericana son la Prueba de Colores y Palabras (Stroop Test), la Prueba de Clasificación de tarjetas del Wisconsin reducido y la Prueba de Símbolo-Digito. A continuación, se realizará una breve descripción de cada una de ellas.

Actualmente se utilizan en Latinoamérica, en general, y en México en lo particular pruebas neuropsicológicas que sean diseñado, desarrollado y estandarizado en otros países como lo son por mencionar unos ejemplos la Figura Compleja de Rey-Osterrieth, desarrollada en Francia y baremada en España, la Prueba de Colores y Palabras de Stroop, diseñada originalmente en 1935 por Stroop en Estados Unidos de Norteamérica (Stroop, 1935), Las Pruebas de Fluidez Verbal, la Prueba de Denominación de Boston, la Prueba de Clasificación de Tarjetas de Wisconsin en sus versiones originales, Corta y Modificada, y así un largo etcétera.

Esta condición de utilizar pruebas que no poseen datos normativos para la población a la que pertenece la persona que evaluamos nos coloca en el predicamento de comparar su

ejecución o sus puntajes contra datos extraídos de poblaciones con otras características sociales, culturales, con otros usos y costumbres, con otras especificidades en el manejo del lenguaje, entre otros, que no son las propias de la persona evaluada. Este hecho conlleva la alta probabilidad de cometer errores en la toma de decisiones al momento de hacer el diagnóstico y las estrategias de trabajo para los pacientes.

## **2.2 Objetivo general**

Generar datos normativos de tres pruebas neuropsicológicas que evalúan de manera general funciones como la atención, el procesamiento de información y funciones ejecutivas en la población de Mexicali, Baja California.

## **2.3 Objetivos específicos**

Generar datos normativos para la Prueba de Colores y Palabras de Stroop en personas sanas de la ciudad de Mexicali en Baja California.

Generar datos normativos para la Prueba de Símbolo-Dígito en personas sanas de la ciudad de Mexicali en Baja California.

Generar datos normativos para la Prueba de Modificada de Clasificación de Tarjetas de Wisconsin en personas sanas de la ciudad de Mexicali en Baja California.

## **2.4 Datos demográficos de Baja California y Mexicali**

Dado que el objetivo de esta tesis es generar datos normativos de la Prueba de Colores y Palabras de Stroop, la Prueba Modificada de Clasificación de Tarjetas de Wisconsin y la Prueba de Símbolo Dígito en un grupo de personas sanas de la ciudad de Mexicali en Baja California es necesario contextualizar geográficamente y socioculturalmente a la ciudad de Mexicali.

Su localización geográfica es en la frontera noroeste de México y no solamente se debe considerar eso, sino que además está en la frontera con el Estado de California, el más rico de los Estados Unidos de Norteamérica. Este hecho le da a Mexicali un carácter de importancia para la migración de miles de personas que provienen de diferentes estados de la República Mexicana y de distintos países de Centroamérica como Guatemala, Honduras, El Salvador. También se han registrado flujos migratorios de Cuba, Haití y África. Asimismo, desde sus inicios Mexicali ha contado con una gran población proveniente de China, misma que tiene una gran fuerza laboral y poder económico.

Baja California es una de las 32 entidades federativas de México. De 1823 a 1952 era un territorio que dependía económica y políticamente del gobierno central de México. En 1952 se convierte en el Estado 29 con 4 municipios (Mexicali, Tijuana, Ensenada y Tecate) actualmente cuenta con 5 municipios (el más reciente es Playas de Rosarito, en 1995). Su ciudad capital es Mexicali.

Según la Encuesta Intercensal de 2015 (INEGI, 2016) la población de Baja California es de 3,315,766 habitantes. El 50.2 % son mujeres y el 49.8 % son hombres. El 92 % viven en la zona urbana y el 8 % restante en la rural. El promedio de escolaridad en el estado es de 9.8 años completados. A nivel nacional ocupa el lugar 14 en cuanto a su población total. Su población económicamente activa son 786,056 trabajadores, de los

cuales el 59 % son hombres y el 41 % son mujeres. La Tabla 1 muestra las principales características demográficas de Baja California de acuerdo al INEGI (2016).

Tabla 1

Indicadores Demográficos de Baja California, Fuente: Consejo Nacional de Población.

Indicador	
Población a mitad de año	3,682,063
Hombres	1,842,355
Mujeres	1,839,708
Nacimientos	61,262
Defunciones	19,983
Crecimiento natural	41,279
Inmigrantes interestatales	41,295
Emigrantes interestatales	35,714
Inmigrantes internacionales	7,976
Emigrantes internacionales	6,948
Migración neta interestatal	5,581
Migración neta internacional	1,028
Tasa bruta de natalidad*	16.64
Tasa bruta de mortalidad*	5.43
Tasa de crecimiento natural**	1.12
Tasa de crecimiento social total**	0.18
Tasa de crecimiento total**	1.30
Tasa global de fecundidad	1.98
Esperanza de vida al nacimiento total	74.69
Esperanza de vida al nacimiento hombres	71.66
Esperanza de vida al nacimiento mujeres	77.88

NOTA: Adaptado de INEGI. \* Por mil \*\* Por cien.

Sus principales sectores de actividad económica son: Primaria (agricultura, ganadería, entre otros): 2.8 %; Secundaria (manufactura, construcción, entre otros): 41.7 % y Terciaria (comercio, restaurantes, servicios entre otros).

## 2.5 Datos demográficos de Mexicali

Mexicali es la capital de Baja California, el Consejo Nacional de Población (CONAPO) estimó en 2017 que su población era de 1,052,656 habitantes, de los cuales 528,857 (50.2 %) eran hombres y 523,799 (49.8 %) eran mujeres. La edad promedio era de 28 años. La

distribución por grupos de edad era la siguiente: 19.9 % son de 0 a 11 años de edad, 31.1% son de 12 a 29, 42.5 % son de 30 a 64 y 6.5 % son de más de 65 años. Respecto de la educación, según la Encuesta Intercensal 2015 realizada por el Instituto Nacional de Estadística, Geografía e Informática (INEGI, 2015) en Mexicali había un total de 729,503 personas con más de 15 años de edad, de los que el 98.1 % saben leer y 1.9 % eran analfabetas. En esta misma encuesta se encontró que en este grupo mayor de 15 años de edad, el 50.6 % habían terminado la educación básica (pre-escolar, primaria y secundaria), el 24.9 % habían terminado la educación media superior (estudios técnicos o comerciales, bachillerato), el 21.9 % habían terminado la educación superior (licenciatura, maestría o doctorado) y el 2.5% no tenía escolaridad. El 84.3 % estaban afiliados a alguno de los institutos de salud del estado y el 15.5 % no lo estaban, menos del 1 % no contestaron la encuesta. La Tabla 2 muestra algunas características demográficas de Mexicali de acuerdo al Comité de Planeación para el Desarrollo del Estado de Baja California (COPLADE, 2107)

Tabla 2

Indicadores Demográficos de Mexicali, Fuente: COPLADE.

Indicador	
Población a mitad de año	1;052,656
Hombres	528,857
Mujeres	523,799
Edad promedio	28 años
Distribución por grupos de edad	
0 a 11 años	19.9%
12 a 29 años	21.1%
30 a 64	42.5%
65 o más años	6.5%
Personas mayores de 15 años de edad	
Saben leer	98.1
No saben leer	1.9%

Tabla de elaboración propia



Figura 6. Representación de la ubicación Geográfica de Mexicali, Baja California.

## 2.6 Justificación

Lo que se ha revisado hasta el momento con respecto a cada una de las 3 pruebas neuropsicológicas y la presentación de los datos demográficos subrayan la importancia de tener datos normativos para cada una de las poblaciones con las que trabajemos dado que cada localidad, estado, región o país tiene especificidades que los hacen únicos. Es decir, que las características de edad, de escolaridad de cada persona evaluada o la escolaridad de los padres en el caso de niños, el sexo y sobre todo las interrelaciones que puedan generarse entre estas variables, determinarán la manera en la que cada persona responderá en la evaluación.

Además de que cada prueba que se utilice en la evaluación debe de reunir ciertos criterios como confiabilidad, validez, datos normativos, entre otros. Es deseable que estos datos normativos sean obtenidos tomando en cuenta todas las variables que pueden influir en la ejecución de la persona. La mayoría de los datos normativos generados para las pruebas psicométricas y neuropsicológicas se han obtenido con métodos en los que se han convertido las puntuaciones brutas a escalas como valores Z o puntajes T, o se obtienen las medias y desviaciones estándar para cada variable y grupo que se evalúa. Sin embargo, cuando se obtienen datos normativos con estas metodologías se pueden presentar las siguientes limitaciones: 1) se asume que algunas variables tienen relevancia con el proceso evaluado y no se comprueba si así es. Por ejemplo, la edad, el sexo o el nivel de escolaridad de la persona se consideran como relevantes en la ejecución de las pruebas, sin embargo, en la mayoría de los reportes publicados no se han realizado pruebas para saber el peso de cada una de estas variables o de sus interrelaciones; 2) el tamaño de la muestra de la que se obtienen los datos se reduce a medida que se dividen los grupos por edad, por sexo y por escolaridad. Por ejemplo, si se quisiera obtener datos normativos ajustados a dos variables, como el sexo y la edad; al dividir la muestra en dos sexos, su tamaño disminuye en aproximadamente un 50%

y si se dividiera en 5 grupos de edad, el tamaño disminuirá hasta en un 10%. Si adicionalmente se consideraran grupos de años de escolaridad, la muestra disminuiría todavía más. Al dividir la muestra en subgrupos se pueden obtener datos normativos estimados de manera menos precisa, ya que la discrepancia entre los estadísticos de la muestra y los parámetros de la población real se incrementan a medida que disminuye el tamaño de la muestra. Adicionalmente, cuando los datos normativos se dividen por grupos de edad, los límites superiores e inferiores pueden generar errores en la toma de decisiones. Por ejemplo si los datos normativos se dividieran en grupos de edad con rangos de 2 años (4-5, 6-7, 8-9, etc.) y se evaluara a un niño de 8 años y 1 mes de edad y a otro niño con 7 años y 11 meses de edad a cada uno se le compararía con datos normativos diferentes cuando la diferencia es de un par de meses y; por otra parte, si se evaluara a dos niños, uno con 6 años y 1 mes y el otro con 7 años y 11 meses ambos se compararían con el mismo dato normativo, pero la diferencia de edades entre ambos es de casi dos años. Claro que este problema se puede resolver elaborando tablas de datos normativos con rangos de edades más reducidos, pero eso implicaría que la muestra total debe de dividirse en más subgrupos, con el problema de reducir el tamaño de cada subgrupo; otra opción sería aumentar significativamente el tamaño de la muestra, con los consecuentes problemas de logística, y aumento de gastos. Recientemente se han realizado trabajos para desarrollar datos normativos con una técnica estadística que soluciona los problemas de utilizar puntajes Z o T, o de utilizar datos basados en medias y desviaciones estándar. Esa técnica es el análisis de regresión múltiple.

El objetivo de esta tesis es generar datos normativos de la Prueba de Colores y Palabras de Stroop, la Prueba Modificada de Clasificación de Tarjetas de Wisconsin y la Prueba Símbolo Dígito en un grupo de personas sanas adultas de la ciudad de Mexicali en Baja California, utilizando modelos de regresión múltiple.



### Capítulo 3. Método

#### 3.1 Método

##### 3.1.1 Participantes y tamaño de la muestra.

Este trabajo hace parte de un estudio multicéntrico que tenía como propósito generar datos normativos para población adulta de habla hispana de Latinoamérica (Guàrdia-Olmos, Perú-Cebollero, Rivera, y Arango-Lasprilla, 2015). Para el caso de México, se recolectaron datos en cuatro ciudades (México D.F., Guadalajara, Monterrey y Mexicali), usando un muestreo no probabilístico por cuotas, donde se tomó cada ciudad como una cuota. En la presente tesis doctoral se analizaron los datos correspondientes a la ciudad de Mexicali, donde la muestra inicial estuvo conformada por 453 adultos clínicamente sanos, residentes en la ciudad de Mexicali, Baja California, México; sin embargo, se eliminaron 23 participantes por dejar inconclusas las pruebas estudiadas. La muestra final estuvo conformada por 430 participantes.

La edad promedio de la muestra final fue de 50.6 años ( $DE=19.2$ ) y un nivel escolar promedio de 9.8 años ( $DE=5.0$ ). La mayoría de la muestra (70.5%) eran mujeres. Del total de la muestra el 85.9 % residían en el área urbana. La información demográfica de la muestra se encuentra con mayor detalle en la Tabla 3.

##### 3.1.2 Criterios de inclusión y exclusión.

Los criterios de inclusión fueron: a) Tener entre 18 y 90 años de edad, b) haber nacido y residir en el momento de la evaluación en la ciudad de Mexicali, Baja California, c) tener el español como lengua materna, d) no tener una ocupación relacionada o familiarizada con el uso y/o manipulación de las pruebas incluidas en el protocolo de investigación, e) ser capaz de leer y escribir, f) obtener al menos 23 puntos en el Mini Examen del Estado Mental (Folstein, Folstein y McHugh, 1975; Villaseñor-Cabrera et al, 2010), g) obtener menos de 5

puntos en el Cuestionario de Salud del Paciente (Kroenke, Spitzer y Williams, 2001) y h) obtener al menos 90 puntos en el Índice de Barthel (Mahoney y Barthel, 1965)

Tabla 3  
Características sociodemográficas de la muestra.

Rango de edad	<i>n</i>	Edad	Escolaridad	Sexo	
		$\bar{M}$	$\bar{M}$	Mujeres	Hombres
18 a 22	66	19.3	11.5	42	24
23 a 27	17	25.2	13.2	8	9
28 a 32	15	29.3	11.7	8	7
33 a 37	17	35.3	8.8	14	3
38 a 42	18	39.8	7.9	16	2
43 a 47	24	44.8	10.3	12	12
48 a 52	34	49.9	9.2	26	8
53 a 57	61	55.4	9.7	43	18
58 a 62	57	60.1	9.7	45	12
63 a 67	40	65.1	8.5	32	8
68 a 72	27	69.6	10.7	21	7
73 a 77	22	74.6	8.5	15	7
78 a 82	19	79.6	7.7	11	8
82 o más	12	86.2	5.9	10	2
<b>Total</b>		<b>50.6</b>	<b>9.8</b>	<b>303</b>	<b>127</b>

*n*= tamaño de la muestra,  $\bar{M}$ = media, *DE*= desviación estándar

Se excluyeron a aquellos participantes que tuvieron al menos una de las siguientes características: a) haber presentado alguna enfermedad del sistema nervioso central, con la presencia de déficits neuropsicológicos, en algún momento de su vida; b) estar diagnosticado con una enfermedad sistémica asociada con deterioro cognitivo (ej. diabetes mellitus, hipotiroidismo); c) tener historial de abuso de alcohol u otras sustancias psicoactivas; d) presentar antecedentes de enfermedad psiquiátrica; e) presentar déficits sensoriales visuales y/o auditivos que afectaran la administración de las pruebas; f) consumir medicamentos psiquiátricos o de otro tipo que afectara el rendimiento cognitivo; g) tener una historia de consumo de medicamentos contra el dolor crónico (ej. inhibidores de la monoamino oxidasa);

h) no cubrir el puntaje requerido en cualquiera de las pruebas de Screening (Mini Mental, Cuestionario de Salud o Índice de Barthel).

Los criterios de inclusión y exclusión se revisaron junto con una entrevista inicial y con la aplicación de las Pruebas de Screening. Con estos instrumentos se revisaban las características de los participantes que permitieran decidir si se continuaba con la evaluación neuropsicológica o si se excluían del estudio. Si alguno de los participantes no obtenía el puntaje mínimo requerido en cualquiera de las pruebas de screening, se excluía de la muestra y no era evaluado. La entrevista incluía el llenado de un formato con los datos demográficos de cada participante con el propósito de poder asignarlo a los diferentes grupos de edad y de escolaridad que se analizaron estadísticamente.

El tamaño de muestra total se estableció mediante la estimación clásica a partir de los supuestos de población infinita y bajo el caso de la máxima indeterminación ( $\pi = 1 - \pi = 0.5$ ). A la vista de los valores obtenidos se optó por un tamaño de muestra de 430 participantes, lo que supone un nivel de confianza del 95 % con un error máximo del 4.5% aproximadamente (Guàrdia-Olmos et al, 2007). La recolección de la muestra se realizó a través de un muestreo no probabilístico, accidental y de conveniencia (Fontes de Gracia et al, 2018), donde la muestra fue restringida a la parte de la población que fue fácilmente accesible, teniendo en cuenta el azar mas no la aleatoriedad (Ospina Botero, 2001).

### **3.2 Instrumentos**

#### **3.2.1 Entrevista inicial.**

Con esta entrevista se recogió información relacionada con los datos sociodemográficas del paciente tales como: identificación del participante, fecha de la aplicación del protocolo, datos demográficos, laborales y económicos, estado de salud

general y los tratamientos farmacológicos actuales. Esto permitió hacer la primera evaluación con el propósito de saber si la persona cumplía con los criterios de inclusión. De igual forma se obtenía información sociodemográfica para agregar al participante en el grupo de edad y escolaridad que le correspondía.

### 3.2.2 Pruebas de screening.

En este estudio se utilizaron lo que se nombró Pruebas de Screening, que son tres pruebas que se aplicaban a cada participante y en las que debían de obtener un puntaje determinado para continuar con la aplicación de las pruebas neuropsicológicas. Estas pruebas aseguraban que cada participante tuviera un mínimo desempeño cognoscitivo general adecuado para completar las pruebas neuropsicológicas, también daban información si la persona se encontraba deprimida y finalmente nos aseguraban que el participante no dependiera físicamente de otras personas.

Los participantes tenían que cubrir el criterio de ejecución en las tres pruebas de screening para poder ser incluidos en la muestra. A continuación, se describen estas pruebas.

#### 3.2.2.1 *Mini examen del estado mental. Mini-mental state examination (MMSE).*

En 1975 se publicó el primer artículo sobre el Mini-Mental State Examination (MMSE, Folstein, Folstein y McHugh, 1975). Desde su publicación, se convirtió en un instrumento de evaluación cognoscitiva en el área clínica, sin embargo, el uso del MMSE ha extendido su aplicación a la docencia y a la investigación, concretamente con pacientes en edad geriátrica. La prueba es una medición breve y cuantitativa que puede utilizarse para detectar déficit cognoscitivo, para evaluar la gravedad en un determinado momento, para seguir el curso de los cambios cognoscitivos en el tiempo y para evaluar la respuesta a un tratamiento. Además, es un test fácil de utilizar y es bien aceptado tanto por los evaluadores como por los evaluados.

Diversas investigaciones han demostrado que el MMSE es un instrumento fiable y válido en distintas muestras de pacientes y población en general. Por ejemplo, Khon, Vicente, Rioseco, Saldivia y Torres (2008) aplicaron esta prueba a 2978 chilenos con edades mínimas de 15 años y encontraron relaciones entre el puntaje obtenido y el estado civil, el ingreso y la zona que habitaba la persona (rural o urbana). En otro estudio, Raina, Chander, Raina y Grover (2016), aplicaron el MMSE a una muestra de analfabetos en el Noroeste de la India y realizaron un análisis de componentes principales, encontrando que las variables a considerar al evaluar a personas analfabetas de La India son la orientación espacio-temporal y la memoria inmediata. En un estudio realizado en México (Villaseñor-Cabrera, Guardia-Olmos, Jiménez-Maldonado, Rizo-Curiel y Perú-Cebollero, 2010) se determinó que la sensibilidad y especificidad del MMSE era buena, aunque en la población mexicana se debería de agregar un puntaje de corrección al considerar el efecto de la edad y la escolaridad y ellos propusieron un punto de corte entre los 24 y 25 puntos. Por último, el Mini-Mental, se ha convertido, en el campo de los trastornos cognoscitivos, en el test de “screening” más utilizado en la clínica, en la investigación epidemiológica internacional y en diversos ensayos clínicos en los que se debe realizar evaluaciones de las funciones cognoscitivas de los pacientes.

#### 3.2.2.2 Cuestionario de salud del paciente. *Patient health questionnaire (PHQ-9)*.

Se trata de un breve cuestionario que evalúa la presencia de trastorno depresivo mayor basándose en los criterios del DSM-IV-R (Kroenke, Spitzer y Williams, 2001; Diez-Quevedo, Rangil, Sanchez-Planell, Kroenke y Spitzer, 2001). El PHQ-9 es un cuestionario que evalúa 9 criterios de depresión según el DSM. El PHQ-9 ha demostrado ser útil como herramienta para el diagnóstico de los síntomas de la depresión, con confiabilidad, validez, sensibilidad y especificidad aceptables (Cassiani-Miranda, Pérez-Anibal, Vargas-Hernández,

Herazo-Bustos y Cabarcas-Tovar, 2018; González-Blanch, et al, 2018; Hancock y Larner, 2009; Kohrt, Luitel, Acharya y Jordans, 2018). Además de su uso como instrumento de diagnóstico, el PHQ-9 se puede también utilizar como herramienta para determinar la severidad de la depresión para supervisar su tratamiento (Hancock y Larner, 2009). Concretamente, el PHQ-9 ha demostrado una sensibilidad del 88.2% y una especificidad del 86.6% en la detección de pacientes depresivos, al establecer comparación con la escala de depresión de Beck (Urtasun et al, 2019). Así mismo, tiene una validez de constructo y una validez predictiva concurrente con los criterios CIE-10 para la depresión. De igual manera, se ha demostrado una consistencia interna medida por el coeficiente de Cronbach, obteniendo una puntuación de .87 (Urtasun et al, 2019). El PHQ-9 se ha utilizado en muestras de habla hispana mexicana (Arrieta et al, 2017), colombiana (Arango-Lasprilla y Rivera, 2015; Cassiani-Miranda, Pérez-Anibal, Vargas-Hernández, Herazo-Bustos y Cabarcas-Tovar, 2018), argentina (Urtasun et al, 2019) y española (Diez-Quevedo, Rangil, Sanchez-Planell, Kroenke y Spitzer. 2001). Inclusive se ha utilizado en estudios multicéntricos con muestras de 11 países latinoamericanos (Guardia-Olmos, Perú-Cebollero, Rivera y Arango-Lasprilla, 2015)

### *3.2.2.3 Índice de Barthel.*

El índice fue diseñado por Mahoney y Barthel en 1955 (Mahoney y Barthel, 1965). Originalmente fue creado para la medición de la evolución de sujetos con procesos neuromusculares y musculo-esqueléticos. Se trata de un índice que consta de diez parámetros y mide las actividades básicas de la vida diaria (ABVD). Además, se trata de uno de los instrumentos para medir ABVD más utilizado internacionalmente, especialmente en pacientes con enfermedades cerebrovasculares agudas (Lee, Kim, Sohn, Lee, Lee, Shin, et al., 2020).

La reproducibilidad inter e intra-observador ha sido demostrada en diversos estudios con coeficientes de correlación de entre .88 y .98 (Lee, Kim, Sohn, Lee, Lee, Shin, et al., 2020; Unnanuntana, Jarusriwanna y Nepal, 2018). Referente a la validez, se trata de un índice que predice bien la mortalidad (Yalcinli, Ersel, Karbek-Akarca, Can y Midik, 2015) y la respuesta a los tratamientos de rehabilitación (Unnanuntana, Jarusriwanna y Nepal, 2018). Además, valora con precisión la duración de la estancia hospitalaria, la necesidad de hospitalización, el resultado funcional final y la capacidad para seguir viviendo en la comunidad. Así mismo, existe una buena relación entre la evaluación realizada al alta y el índice de Barthel.

### 3.2.3 Pruebas neuropsicológicas.

#### 3.2.3.1 Prueba de símbolo y dígitos (Smith, 2002).

La Prueba de símbolos y dígitos evalúa la velocidad de procesamiento, el rastreo visual y el seguimiento visual. Consta de una clave compuesta por 9 símbolos abstractos y a cada símbolo le corresponde un número del 1 al 9. La tarea de la persona evaluada es escribir el número que le corresponde a cada símbolo en una hoja de respuestas en la que aparece solo el símbolo. La persona tiene que hacer un ensayo de nueve reactivos y posterior a esto se le indica que deberá de escribir los números sin saltar ninguna celda. Se le dan 90 segundos para que realice la tarea y al final se cuentan las respuestas correctas. El máximo puntaje es 110.

#### 3.2.3.2 Prueba de colores y palabras de Stroop (Golden, 2010).

Esta prueba evalúa la habilidad que tiene una persona para suprimir una respuesta habitual por otra menos familiar. También evalúa la flexibilidad cognoscitiva, la atención selectiva y funciones ejecutivas. Esta versión de la prueba está compuesta de tres hojas en las que se presentan 100 estímulos organizados en 5 columnas con 20 reactivos cada una. En

la primera hoja se presentan escritos los nombres de tres colores (azul, verde y rojo) impresos en tinta negra y la persona debe leerlos lo más rápido y claro que pueda. Una vez iniciada la tarea tendrá 45 segundos para leer y al término de este tiempo se le pide que se detenga. En la segunda hoja se presentan 100 grupos de 4 letras X (XXXX), igualmente organizadas en 5 columnas. Cada grupo de letras está impreso en tinta roja, azul o verde y la persona debe decir el color en el que está impreso cada grupo de letras X. También se le dan 45 segundos para que realice la tarea. En la tercera y última hoja se le presentan las mismas palabras que se presentaron en la primera hoja pero ahora cada palabra (nombre de un color) está impreso en una color de tinta diferente al color escrito (por ejemplo, la palabra ROJO está impresa en tinta de color AZUL), y la persona deberá de decir el color en el que está impresa la palabra (DECIR AZUL) y debe de inhibir la lectura de esa palabra ( NO LEER la palabra ROJO). También en esta parte de la prueba se le dan 45 segundos. Al final se calcula un Índice de Interferencia con la siguiente fórmula:  $PC - [(Px C)/(P + C)]$ . Con esto se obtiene un indicador del grado de control que tiene la persona sobre la interferencia

### *3.2.3.3 Prueba modificada de clasificación de tarjetas de Wisconsin (Schretlen, 2010).*

La prueba consta de 4 tarjetas clave o guía y de 48 tarjetas que deberán de clasificarse de acuerdo a criterios como el color, la forma o el número de estímulos que aparezcan en cada tarjeta. A la persona evaluada se le pide que clasifique cada en base a las 4 tarjetas clave (en una aparece un triángulo rojo, en la siguiente dos estrellas verdes, en la tercera tres cruces amarillas y en la última cuatro círculos azules) deberá clasificar cada una de las 48 tarjetas. Deberá hacerlo tomando una tarjeta a la vez, y una vez hecho la clasificación de la tarjeta se le deberá decir en voz alta si su elección fue CORRECTA o INCORRECTA. También se le



mencionar que la manera de clasificar las tarjetas puede cambiar y que cuando eso ocurra, se le avisará para que cambie de criterio de clasificación.

### **3.3 Procedimiento**

#### **3.3.1 Diseño de investigación utilizado.**

Estudio correlacional explicativo, con un diseño no experimental de tipo transversal de acuerdo con Hernández-Sampieri, Fernández y Baptista (2003). Es un estudio correlacional explicativo dado que se miden algunas variables como la edad, la escolaridad y el sexo, además trata de explicar cómo se relacionan con la ejecución cognoscitiva en tres pruebas neuropsicológicas y también cómo es qué ese tipo de relaciones afectan la ocurrencia y el grado de ocurrencia de estas variables; y es un diseño no experimental de tipo transversal porque no se manipulan deliberadamente variables, y la medición se hace solo una vez.

#### **3.3.2 Recolección de datos.**

El presente estudio forma parte de un proyecto multicéntrico que tiene como objetivo generar datos normativos en pruebas neuropsicológicas en población adulta de habla hispana (Guàrdia-Olmos et al., 2015). El primer paso fue redactar un documento con la propuesta del estudio, la cual incluía metodología y aspectos éticos, y que posteriormente se entregó al Comité de Ética de la Universidad de Deusto (Bilbao, España), como centro responsable del estudio multicéntrico. Después de la aprobación del Comité de Ética, se procedió a comprar los manuales, las hojas de respuestas, y el material (cuadernillos y tarjetas estímulo) de cada una de las pruebas neuropsicológicas a las editoriales cuando era necesario. Seguidamente, las pruebas fueron revisadas por dos especialistas en neuropsicología, nativos de México, con el objetivo de asegurar que el lenguaje de las instrucciones fuera adecuado para la

población mexicana. Adicionalmente, los participantes debían de pasar 3 pruebas de tamizaje o screening y cumplir con los criterios de inclusión y exclusión de este estudio.

Los evaluadores encargados de recolectar la muestra fueron entrenados usando una serie de herramientas y ayudas audiovisuales con el propósito de lograr un proceso estándar de administración de la batería de pruebas neuropsicológicas. Dichas herramientas constaban de: a) una lista aleatorizada para determinar el orden de administración de las pruebas para cada participante, teniendo como objetivo evitar el sesgo de orden y condicionamiento cognitivo (Fontes de Gracia et al, 2018). Para la creación de la lista, se usó la función  $f_x = ALEATORIO()$  en Microsoft Excel© y se configuró teniendo en cuenta la interacción de las pruebas de lenguaje con la prueba de memoria verbal; b) un esquema para la toma de decisiones en el proceso de evaluación de una persona; c) una plantilla en Microsoft Excel© para el ingreso de la información recolectada de cada persona evaluada. La plantilla fue diseñada usando las opciones de configuración: *validación de datos=personalizada* (variables numéricas), *listas desplegables* (variables categóricas) y *formatos de configuración*. La configuración anterior asegura el control del sesgo de ingreso de la información; y d) una serie de ejemplos mostrando los errores más frecuentes en el momento de la administración y calificación de cada una de las pruebas. Los participantes respondieron una batería de diez pruebas neuropsicológicas, sin embargo, en este trabajo solo se reportan los resultados de la Prueba de Símbolo Dígito (Smith, 2002), la Prueba de Colores y Palabras de Stroop (Golden, 2010) y la Prueba Modificada de Clasificación de Tarjetas de Wisconsin (Schretlen, 2010). La recolección de la información inició en marzo de 2013 y finalizó en abril de 2015.

### 3.3.3. Análisis de datos.

#### 3.3.3.1 *Análisis exploratorios de los datos.*

La muestra fue dividida en dos sub-muestras de manera aleatoria. La primera sub-muestra estuvo conformada por el 80% del total de la muestra y la segunda por el 20% restante. La sub-muestra 1 (80% de los datos;  $n= 350$ ) fue usada para generar los datos normativos, mientras que la sub-muestra 2 (20% de los datos;  $n= 80$ ) se usó para comprobar los modelos generados en la sub-muestra 1. En la fase de análisis exploratorio de los datos, se estimó la media y desviación típica para cada una de las puntuaciones estudiadas (Prueba de Colores y Palabras Stroop, Prueba Modificada de Clasificación de Tarjetas de Wisconsin, Prueba de Símbolo Dígitos). Para superar la limitante del muestreo no probabilístico se realizó un análisis de datos para comprobar la aleatoriedad de la secuencia de datos por medio de la prueba de rachas de Wald y Wolfowitz (1943) tomando como punto de corte la mediana.

Para estimar la evidencia de validez de constructo se realizó un análisis factorial confirmatorio asumiendo una estructura a priori de dos factores (factor STROOP: Palabras, Colores y PC; factor M-WCST: Categorías, Errores y Perseveraciones). En este análisis no fue posible incluir la puntuación del SDMT debido a que solamente se obtiene una puntuación única, lo que impide construir un factor (variable latente). Además de lo anterior, para cada una de las pruebas neuropsicológicas se estimó la validez convergente usando estimaciones del coeficiente de correlación de Spearman ( $\rho_{xj}$ ) y el coeficiente de determinación ( $r^2$ ) entre la puntuación de cada prueba y la puntuación total del MMSE. Con respecto a la fiabilidad de las pruebas, se evaluó la consistencia interna a través del alfa de Cronbach (1951) y del coeficiente de confiabilidad del constructo de McDonald (1999).

Finalmente se estimaron los coeficientes de correlaciones entre las puntuaciones neuropsicológicas estudiadas y las variables sociodemográficas (edad, escolaridad y sexo) usando el coeficiente de correlación de Spearman. La exploración se realizó usando la sub-muestra 1 ( $n=350$ ).

### 3.3.3.2 Efectos de las variables demográficas y la derivación de datos normativos.

Las puntuaciones de cada Prueba se analizaron por separado usando la sub-muestra 1. Los efectos de las variables demográficas en las puntuaciones se evaluaron mediante análisis de regresión múltiple (Rivera et al, 2019; 2020). Se estimó un modelo de regresión múltiple para cada puntuación neuropsicológica estudiada, en el que se incluyeron como variables predictoras la edad, edad<sup>2</sup>, escolaridad, escolaridad<sup>2</sup> y el sexo, así como las interacciones con estas variables. Las variables edad y escolaridad se centralizaron antes de calcular su valor cuadrático y sus interacciones, para controlar la multicolinealidad (Kutner, Nachtsheim, Neter, y Li, 2005). La centralización de estas variables se realizó restando el valor de la media muestral tanto de la edad ( $Edad - \bar{X}_{Edad}$ ), como para la escolaridad ( $Escolaridad - \bar{X}_{Esc.}$ ). El sexo se codificó como 1 para hombre y 0 para mujer.

El modelo principal fue:  $y_i = B_0 + B_1 \cdot (Edad - \bar{x}_{Edad})_i + B_2 \cdot (Edad - \bar{x}_{Edad})_i^2 + B_3 \cdot (Escolaridad - \bar{x}_{Esc.})_i + B_4 \cdot (Escolaridad - \bar{x}_{Esc.})_i^2 + B_5 \cdot Sexo_i + B_k \cdot Interacciones_i + \varepsilon_i$ , donde el subíndice  $i$  hace referencia a los participante ( $i = \{1, 2, \dots, n_i\}$ , y  $n_i$  el número de participantes de la muestra), el subíndice  $k$  refiere a los parámetros de regresión ( $k = \{1, 2, \dots, K\}$ , y  $K$  = el número de efectos fijos en el modelo, incluida la intersección). Finalmente, el término  $Interacciones_i$  es una notación corta para referirse a todas las interacciones posibles (dos niveles) entre los parámetros de regresión. Aquellas

variables predictoras que no resultaron significativas ( $p > 0.005 \approx 0.05/\text{número de predictores}$ ), fueron eliminadas jerárquicamente del modelo, hasta obtener un modelo de regresión final para cada puntuación estudiada [Prueba de Colores y Palabras Stroop (Palabras, Colores, Palabra-Color, Interferencia), Prueba Modificada de Clasificación de Tarjetas de Wisconsin (Categorías, Perseveraciones, Total de errores), Prueba de Símbolo Dígitos]. En el procedimiento de creación de los modelos finales, no se eliminó ningún predictor siempre que también se incluyera en un predictor de orden superior o interacción en el modelo (Aiken y West, 1991).

Para todos los modelos finales de regresión múltiple, se evaluaron los siguientes supuestos: a) multicolinealidad, que fue evaluada mediante el cálculo del factor de inflación de la varianza (FIV) el cual no debía ser mayor que 10; y el cálculo de los valores de tolerancia de colinealidad que no deben ser mayores a 1, b) homocedasticidad, la cual fue evaluada agrupando los valores predictivos ( $\hat{Y}_i$ ) de los participantes en cuantiles y aplicando la prueba de Levene a los residuos, c) normalidad de los residuos estandarizados, evaluados con la prueba de Kolmogorov-Smirnov e histogramas, y d) la existencia de valores influyentes, que fueron evaluados a través de los valores máximos de distancia de Cook y relacionándolos con una distribución  $F(p, n - p)$  (Kutner et al., 2005; Cook, 1997).

Los datos normativos para la Prueba de Colores y Palabras de Stroop (Palabras, Colores, Palabra-Color, Interferencia), Prueba Modificada de Clasificación de Tarjetas de Wisconsin (Categorías, Perseveraciones, total de errores) y Prueba de Símbolo Dígito se generaron a través de la metodología de modelos de regresión múltiple y la desviación típica de los valores residuales de estos modelos a través de cuatro pasos descritos a continuación:

1. Utilizando el modelo de regresión final que se ha obtenido en el procedimiento anterior, se calcula el valor predictivo de la puntuación ( $\hat{Y}_i$ ) a partir de los parámetros fijos ( $B$ ) establecidos en el modelo final de regresión usando la ecuación:

$$\hat{Y}_i = B_0 + B_1X_{1i} + B_2X_{2i} + \dots + B_kX_{ki}. \quad (1)$$

2. Obtener el valor residual del modelo ( $e_i$ ), el cual se calcula restando el valor predictivo ( $\hat{Y}_i$ ) a la puntuación directa ( $Y_i$ ) de la prueba:

$$e_i = Y_i - \hat{Y}_i. \quad (2)$$

3. Estandarizar el valor residual obtenido ( $z_i$ ). Para hacer esto, es necesario dividir el valor residual ( $e_i$ ) entre la desviación típica residual ( $DT_e$ ) obtenida en el modelo de regresión, usando la siguiente fórmula:

$$z_i = e_i / DT_e. \quad (3)$$

4. Finalmente, usando la función de distribución acumulativa normal estándar, se obtiene el valor de percentil exacto correspondiente al valor  $z_i$  calculado previamente. Para esta conversión de valor  $z_i$  a percentil exacto será necesario tener en cuenta si el supuesto modelo de normalidad de los residuos estandarizados se cumple en la muestra normativa, en tal caso se usará la función de distribución normal. En caso contrario se debe usar la función de distribución acumulativa empírica de los residuos estandarizados (si los residuos estandarizados no se distribuyeron normalmente en la muestra normativa).

### 3.3.3.2 Comparación de modelos de datos normativos.

Usando la sub-muestra 2, se transformaron las puntuaciones directas a valores estandarizados ( $z_i$ ) y a percentiles exactos usando los modelos generados en la sub-muestra 1. Seguidamente, las puntuaciones de la sub-muestra 2 fueron transformadas a puntuación  $z_i$  y a percentiles a partir de los datos normativos publicados en el año 2015 (Rivera y et al.,

2015; Arango-Lasprilla y et al., 2015a y 2015b); es decir, cada puntuación directa de la submuestra 2 fue transformada a puntuaciones  $z_i$  y a percentil usando dos modelos: el primer modelo resultante de este trabajo y el segundo modelo publicado en el 2015.

Con las puntuaciones  $z_i$  obtenidas, se estimó la existencia de diferencias entre los dos métodos usando el estadístico de contraste de los rangos signados de Wilcoxon, ya que cada puntuación directa fue convertida a valores  $z_i$  por dos modelos diferentes. Todos los cálculos estadísticos se realizaron usando el programa SPSS versión 24 para Windows y R versión 3.5.3.

#### 3.3.4 Aspectos Éticos

Todos los participantes firmaron un consentimiento informado en el que se indicaban las condiciones en que se llevaría a cabo la evaluación, los derechos que tenían como participantes y además se les solicitaba su permiso para la publicación de los resultados del estudio. Se les aseguraba la confidencialidad de los datos y el anonimato de cada uno de los participantes.

## Capítulo 4. Resultados

### 4.1 Resultados

#### 4.1.1 Análisis exploratorios de los datos.

Para este apartado se usó la submuestra 1 (la cual está conformada por el 80%). La edad promedio de esta submuestra fue de 50.3 ( $DE=19.1$ ) y de la educación fue de 9.7 ( $DE=5.1$ ). Los valores promedios, sus desviaciones estándar y medidas de posición de las puntuaciones en la Prueba de Símbolo dígitos (Puntuación total), Prueba de Colores y Palabras Stroop (Colores, Palabras, Palabra-Color e Interferencia) y Prueba Modificada de Clasificación de Tarjetas de Wisconsin (Categorías, Perseveraciones y total de errores) se pueden encontrar en la Tabla 4.

Tabla 4.  
Medidas de tendencia central y de posición de cada una de las puntuaciones estudiadas

<b>Puntuaciones</b>	<b><math>\bar{M}</math></b>	<b><math>DE</math></b>	<b><math>Min</math></b>	<b><math>Max</math></b>	<b><math>Rango</math></b>	<b><math>Asimetría</math></b>	<b><math>\beta_2</math></b>
Palabras	86.5	21.3	17.0	207.0	190.0	0.19	2.76
Colores	61.7	17.4	2.0	115.0	113.0	0.22	0.60
PC	36.1	16.0	0.0	99.0	99.0	1.00	2.09
Interferencia	0.5	11.4	-31.2	45.2	76.3	0.77	1.92
Categorías	4.3	1.9	0.0	6.0	6.0	-0.69	-0.98
Perseveración	5.5	7.5	0.0	45.0	45.0	2.55	8.64
Errores	13.9	11.6	0.0	47.0	47.0	0.74	-0.51
SDMT	36.8	16.0	3.0	100.0	97.0	0.33	0.26

$\bar{M}$  = Media,  $DE$ = Desviación estándar,  $Min$ = Mínimo,  $Max$ = Máximo,  $\beta_2$ =coeficiente de curtosis

En la Tabla 5 se muestran los valores Z y sus niveles de significancia al realizar la prueba de rachas de Wald y Wolfowitz. De acuerdo con estos resultados la distribución de los resultados en las subpruebas de Palabras, Colores y PC del Stroop, así como la de la



prueba SDMT se distribuyeron aleatoriamente. Las distribuciones de los puntajes de las subpruebas de Categorías, Perseveraciones y Errores del M-WCST no se distribuyeron aleatoriamente. Se debe considerar que en el caso de M-WCST los puntajes que se obtienen en Categorías van desde cero a seis y tal vez eso pudiera tener un efecto sobre la distribución de las puntuaciones y lo mismo pudiera ocurrir con las puntuaciones en las subpruebas de Perseveraciones y Errores debido a que Errores se refiere al total de respuestas erróneas producidas por la persona al contestar la prueba y Perseveraciones se refiere a la cantidad de respuestas perseverativas con respecto de la cantidad de respuestas erróneas; por ejemplo alguien puede lograr cuatro categorías correctas y haber emitido 20 Errores y de esos errores 10 fueron respuestas perseverativas. Pero puede darse el caso de que alguien obtenga cuatro categorías con 20 errores y dos respuestas perseverativas.

Tabla 5.  
Prueba de rachas para cada una de las puntuaciones estudiadas.

	Palabras	Colores	PC	Categorías	Perseveraciones	Errores	SDMT
Valor de prueba	88	62	35	5	3	11	36
Casos < Valor de prueba	174	173	166	146	168	173	169
Casos >= Valor de prueba	176	177	184	204	182	177	181
Casos totales	350	350	350	350	350	350	350
Número de rachas	154	158	165	130	113	119	160
Z	-2.355	-1.925	-1.131	-4.535	-6.725	-6.101	-1.693
Sig.	.019	.054	.258	<.001	<.001	<.001	.090

Al momento de comparar la distribución de la muestra con respecto a la distribución poblacional se encontró una diferencia estadísticamente significativa para la edad ( $M= 41.86$  población vs  $52.51$  muestra;  $DE = 0.693$ ;  $p < 0.001$ ) y el sexo. En esta ultima variable se utilizó la  $\chi^2$  ( $1, N= 695,153$ )= $70.3$ ,  $p = 0.0001$ . Sin embargo, es importante resaltar que la muestra contiene unidades muestrales con información de los diferentes estratos de edad, educación y sexo tal como se puede observar en la tabla 3.

Para la estimación de la evidencia de validez de constructo, se valoró la bondad de ajuste de la estructura factorial propuesta a través de un análisis factorial confirmatorio. La estimación se realizó mediante la técnica de máxima verosimilitud (ML) de acuerdo con las propiedades de los modelos de ecuaciones estructurales – SEM a partir de la siguiente expresión:

$$x = \Lambda_x \xi + \delta.$$

Esta expresión implica que cada variable observada ( $x$ ) se descompone en las nuevas variables latentes o no observables ( $\xi$ ) de acuerdo a las cargas factoriales representadas en la matriz ( $\Lambda_x$ ) compuesta por la carga de cada variable observable ( $\gamma_{ij}$ ) al factor latente asignado. Para evaluar este apartado, se dispuso de los valores habituales de ajuste global y de los valores de cada carga factorial. Los valores de ajuste general que se obtuvieron fueron los siguientes:  $\chi^2 = 39.1$ ; ( $gl = 8$ ;  $p < 0.001$ ); Razón  $\chi^2/gl = 4.8$ ; CFI (Comparative Fit Index) = 0.98; IFI (Incremental Fit Index) = 0.98; GFI (Goodness Fit Index)= 0.97; RMSEA (Root Mean Square Error of Approximation) = 0,091 (IC = 0,079 - 0,103); SRMR (Standardized Root Mean Square Residual) = 0,0457. Se evidencia un modelo adecuado indicando una validez de constructo general (Bentler y Dudgeon, 1996; Schreiber, Nora, Stage, Barlow y

King, 2006) donde los puntajes de los test de este estudio se agrupan de manera tal que se confirma el modelo AFE (Ver Tabla 6 y Figura 7).

Tabla 6.

Estimaciones factoriales para AFC

Factor	Puntuación	Estimación	Error Estándar	Valor z
M-WCST	Categorías	-0.892	0.038 **	-23.505
	Errores	1.020	0.034 **	29.835
	Perseveraciones	0.040	20.061 **	1.607
	Colores	0.913	0.041 **	22.551
STROOP	Palabras	0.748	0.043 **	17.230
	PC	0.799	0.042 **	18.760

*Nota: \*\*<0.001, M-WCST: Prueba Modificada de Clasificación de Tarjetas de Wisconsin, STROOP: Prueba de Palabras y Colores de Stroop.*

Con respecto a los descriptivos residuales (Normalized Residuals), tanto la mediana (<0.001) como la media (-0.29), indican que los residuales están cercanos a cero. Los coeficientes de determinación ( $R^2$ ) de cada una de las variables observadas (R-square Endogenous Variables) son bastantes altos ( $R^2 > 0.56$ ), lo cual señala una evidente comunalidad entre las variables observables y su asociación a los factores propuestos. Con todo ello, se evidencia que existe una adecuada validez de constructo en términos globales, objeto de este apartado específico del estudio.

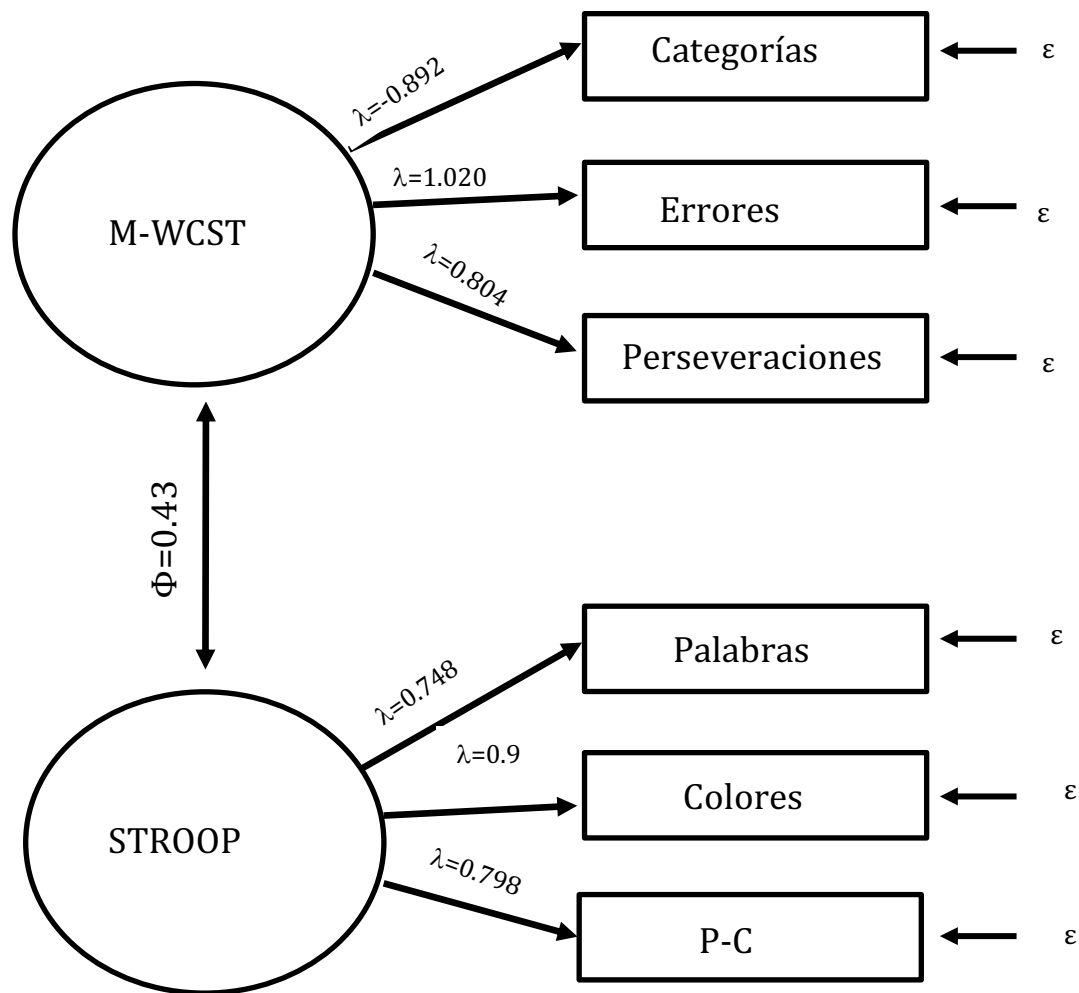


Figura 7. Modelo de Análisis Factorial Confirmatorio

Con respecto a la validez concurrente, tiene el objetivo de establecer si las relaciones observadas entre las puntuaciones de cada uno de las pruebas neuropsicológicas y otras variables externas relevantes son consistentes con la interpretación propuesta para las puntuaciones de las pruebas (Abad et al., 2011). La variable externa con la que se estimaron las correlaciones fue el puntaje directo total del MMSE. El MMSE es una prueba que evalúa el estado cognitivo general (Folstein et al., 1975). Los resultados de este análisis se muestran en la Tabla 7, y se observa que todos los coeficientes de correlación fueron significativos ( $p <$

0.001) y la dirección de las correlaciones fueron consistentes tal y como se esperaba. Por ejemplo, las correlaciones positivas se encontraron entre las puntuaciones del MMSE y todas las puntuaciones del Stroop, en Categorías del M-WCST y en el SDMT.

Tabla 7.

Coeficientes de correlación entre las puntuaciones de cada prueba y el MMSE.

Puntuación	MMSE	r <sup>2</sup>
Stroop Palabras	.332**	0.110
Stroop Colores	.441**	0.194
Stroop PC	.409**	0.167
M-WCST Categorías	.337**	0.038
M-WCST Perseveraciones	-.313**	0.097
M-WCST Errores	-.337**	0.113
SDMT	.441**	0.194

En cuanto a la fiabilidad, se estimaron los coeficientes de consistencia interna Alfa de Cronbach y Omega de McDonald, para las tres puntuaciones de la prueba Stroop, las tres puntuaciones de la M-WCST y finalmente para todas las puntuaciones estudiadas. Se encontraron óptimos coeficientes, tanto para cada una de las pruebas por separado, como para todas las puntuaciones ( $\alpha \geq 0.763$ ;  $\omega \geq 0.861$ ) concluyendo que las puntuaciones tienen una adecuada fiabilidad en términos de consistencia interna (ver Tabla 8).

Tabla 8

Coeficiente de consistencia interna Alfa de Cronbach y Omega de McDonald.

Pruebas	Cronbach's $\alpha$	McDonald's $\omega$
Stroop	0.845	0.861
M-WCST	0.763	0.915
Stroop, M-WCST y SDMT	0.836	0.880

La Tabla 9 muestra las correlaciones entre las puntuaciones estudiadas y las variables demográficas, mostrando alta correlación entre todas las puntuaciones, la edad ( $\rho > |0.285|; p < 0.001$ ) y la educación ( $\rho > |0.379|; p < 0.001$ ). La variable sexo mostró correlación con las puntuaciones de Stroop (Palabra-Color), M-WCST (Categorías, Perseveraciones y Errores) y SDMT ( $\rho > |0.114|; p < 0.05$ ). De manera gráfica se puede observar los coeficientes de correlación en la Figura 8 (puntuaciones SDMT) y Figura 9 (puntuaciones del Stroop).

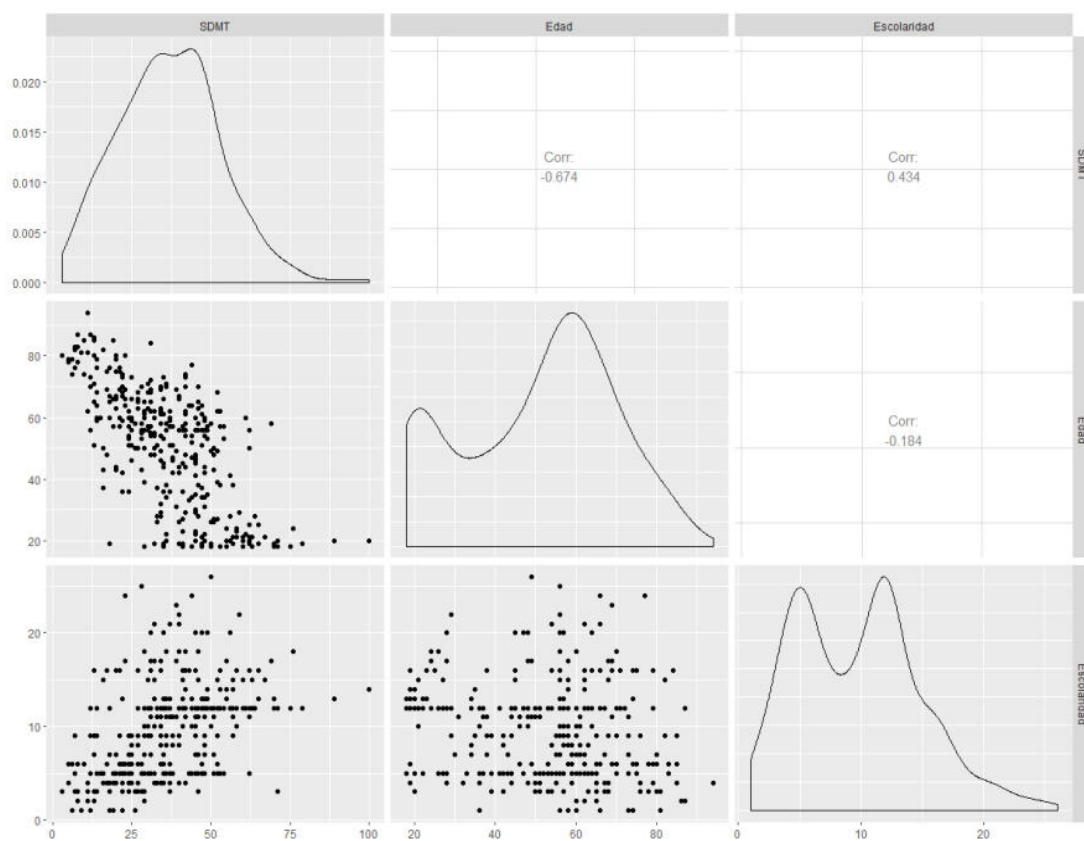


Figura 8. Relación entre las puntuaciones del SDMT y variables demográficas de edad y escolaridad

Tabla 9

Coeficientes de correlación de Spearman

	Edad			Escolaridad			Sexo		
	$\rho$	Intervalo de confianza 95%	$r^2$	$\rho$	Intervalo de confianza 95%	$r^2$	$\rho$	Intervalo de confianza 95%	$r^2$
Palabras	-.413**	(-.496 a -.322)	0.170	.410**	(.319 a .494)	0.168	0.104	(-.001 a .207)	0.010
Colores	-.468**	(-.546 a -.382)	0.219	.406**	(.315 a .490)	0.164	0.082	(-.023 a .185)	0.006
PC	-.491**	(-.567 a -.407)	0.241	.457**	(.370 a .536)	0.208	.120*	(.015 a .222)	0.014
Categorías	-.285**	(-.379 a -.186)	0.081	.379**	(.286 a .465)	0.143	.124*	(.019 a .226)	0.015
Perseveración	.368**	(.274 a .455)	0.135	-.408**	(-.492 a -.317)	0.166	-.136*	(-.237 a -.032)	0.018
Errores	.341**	(.245 a .430)	0.116	-.418**	(-.501 a -.328)	0.174	-.144**	(-.245 a -.040)	0.02
SDMT	-.652**	(-.708 a -.587)	0.425	.475**	(.390 a .552)	0.225	.114*	(.009 a .216)	0.012

Nota: \*\* p&lt;0.001 \* p&lt;0.05.



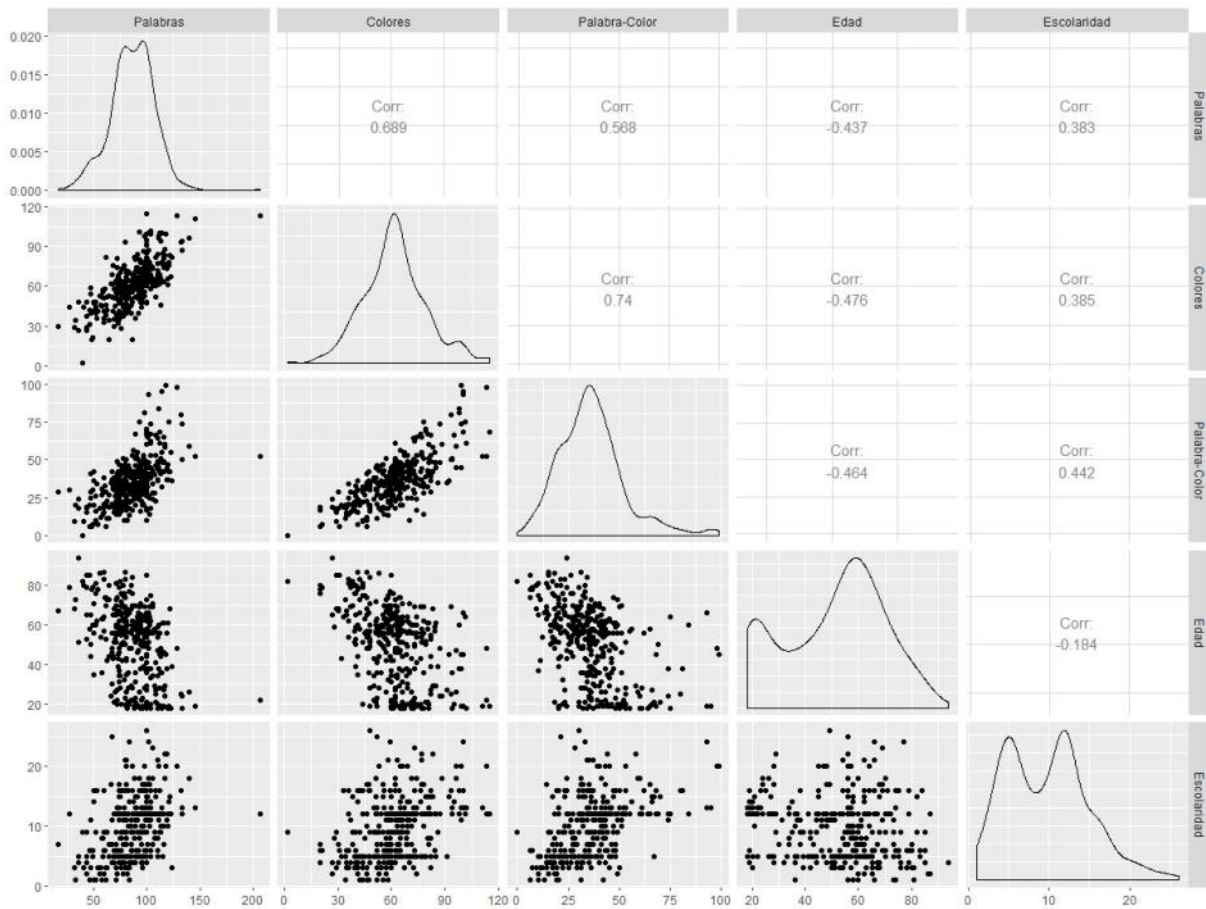


Figura 9. Relación entre las puntuaciones del Stroop y variables demográficas

#### 4.1.1.1 Comprobación de supuestos de los modelos finales.

Los modelos finales de regresión no presentaron multicolinealidad, ya que los valores FIV fueron menores a 2.451 (10 máximo permitido), y los valores de tolerancia de colinealidad fueron menores o iguales a 1 (Ver Tabla 9).

Tabla 9.

Valores de cada uno de los supuestos estadísticos de los modelos finales de regresión en cada puntuación

Puntuaciones	Multicolinealidad		Homocedasticidad	Normalidad	Valores influyentes	
	FIV	Tolerancia	Levene Test	Kolmogorov-Smirnov	Distancia Cook	Percentil
Palabras	1.035	0.966	1.594	0.719	0.106	4%
Colores	1.094	0.966	3.944	0.472	0.055	1%
PC	1.738	0.913	9.701**	1.117	0.261	5%
Interferencia	1.035	0.966	9.034**	0.884	0.123	5%
Categorías	1.035	0.966	16.073**	2.394**	0.037	1%
Perseveraciones	1.094	0.966	21.619**	2.717**	0.152	4%
Errores	1.035	0.966	11.970**	1.782**	0.033	1%
SDMT	2.451	0.848	3.021	0.731	0.167	1%

Nota: \*\*  $p < 0.001$

Los modelos de regresión final para los modelos de las puntuaciones de Palabra, Colores y PC mostraron homogeneidad ( $p > 0.001$ ). Mientras que los modelos finales de regresión para las puntuaciones de Palabra-Color, Interferencia, Categorías, Perseveraciones y total Errores mostraron ser heterogéneos según el test de Levene (ver Figura 10), mostrando una varianza de los errores a lo largo de las observaciones, lo cual tiene un efecto directo en el proceso de estandarización de los residuos. Para corregir la falta de homogeneidad, se dividió en cuatro grupos los valores predictivos  $\hat{Y}_i$  y se estimó una desviación típica residual ( $DE_e$ ) para cada uno de estos cuatro grupos.

Los residuos estandarizados se distribuyeron de manera normal para los modelos de las puntuaciones de la prueba Stroop (palabras, colores, palabra-color, e interferencia) y

SDMT, mientras que para los modelos finales de las puntuaciones del M-WCST (categorías, perseveraciones y total errores), los residuos estandarizados no tuvieron una distribución normal según los valores del test de Kolmogorov-Smirnov y confirmado gráficamente con histogramas (ver Figura 11). Para los residuos estandarizados que no se distribuyeron según la distribución normal, fue necesario crear una distribución empírica acumulada para la conversión a percentiles en el cálculo de datos normativos.

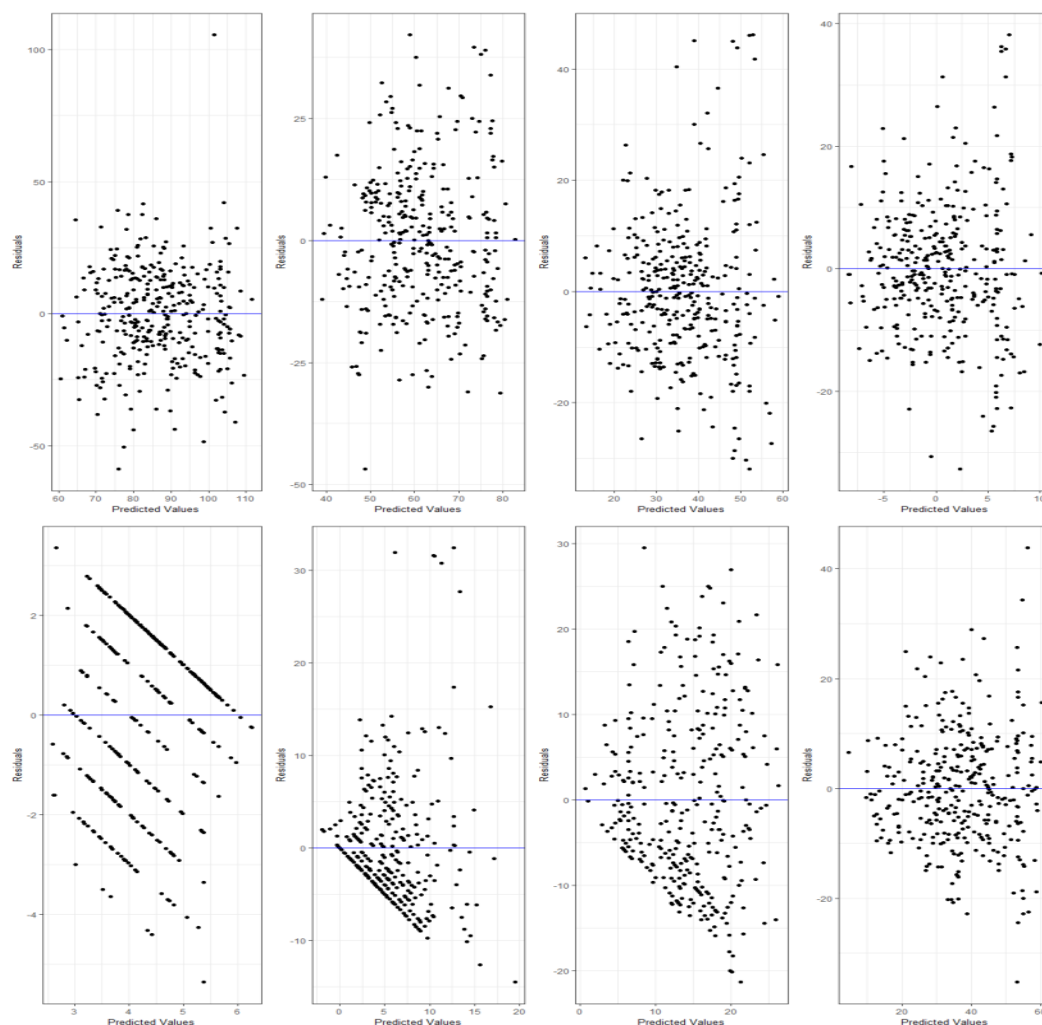


Figura 10. Diagramas de dispersión de los valores predictivos vs. Valores residuales para cada uno de los modelos. De izquierda a derecha arriba: Palabras, Colores, Palabra-Color, Interferencia; abajo: Categorías, Perseveraciones, total Errores y SDMT.

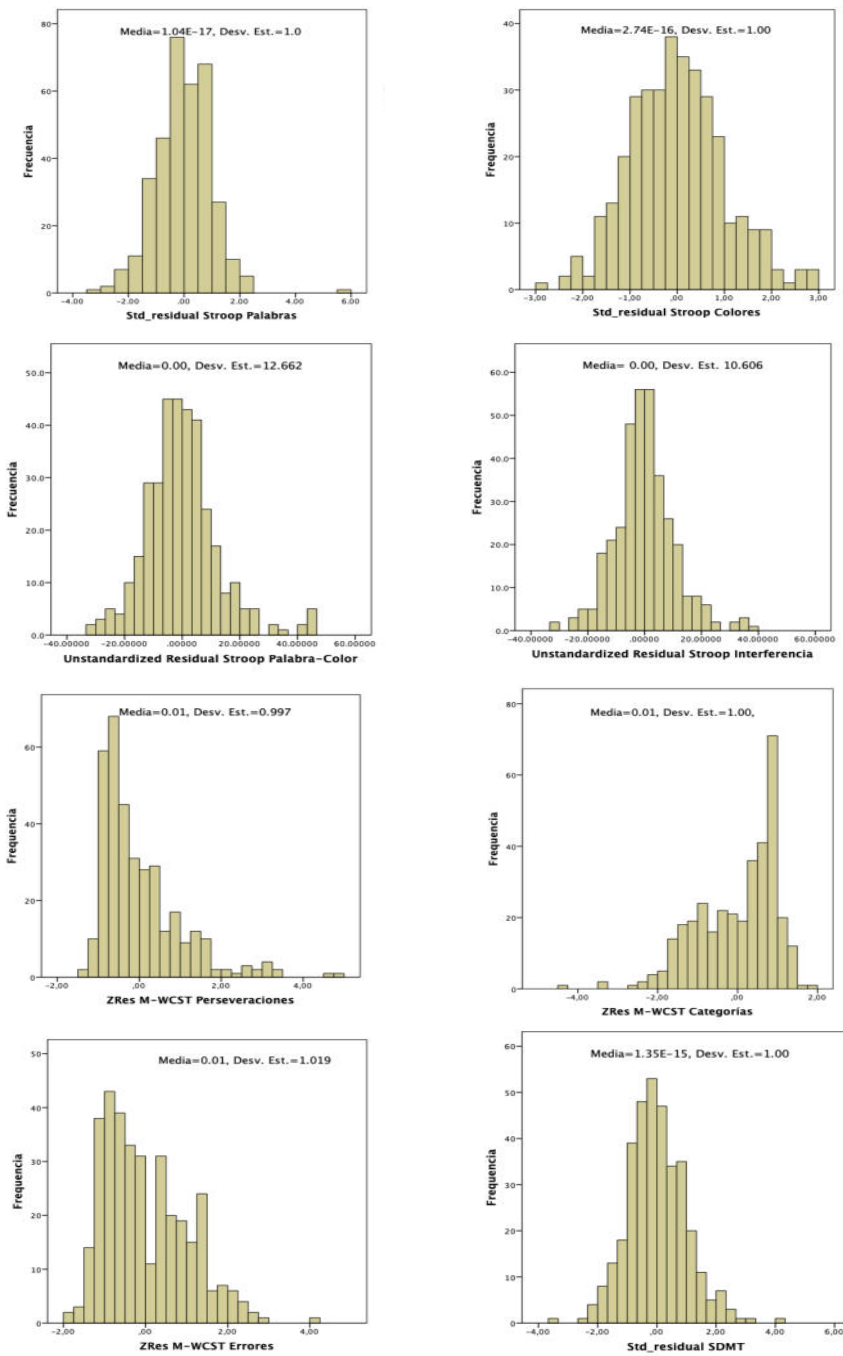


Figura 11. Histogramas de las distribuciones de los residuos estandarizados de cada una de las pruebas aplicadas. De izquierda a derecha y de arriba abajo: Palabras, Colores, Palabra-Color, Interferencia, Categorías, Perseveraciones, total Errores y SDMT.

No se encontraron valores influyentes, siendo la máxima distancia Cook de 0.261 con un percentil de 5, valor muy inferior al punto límite (50). La Figura 12 muestra la distancia Cook de los modelos finales de cada uno de las puntuaciones estudiadas.

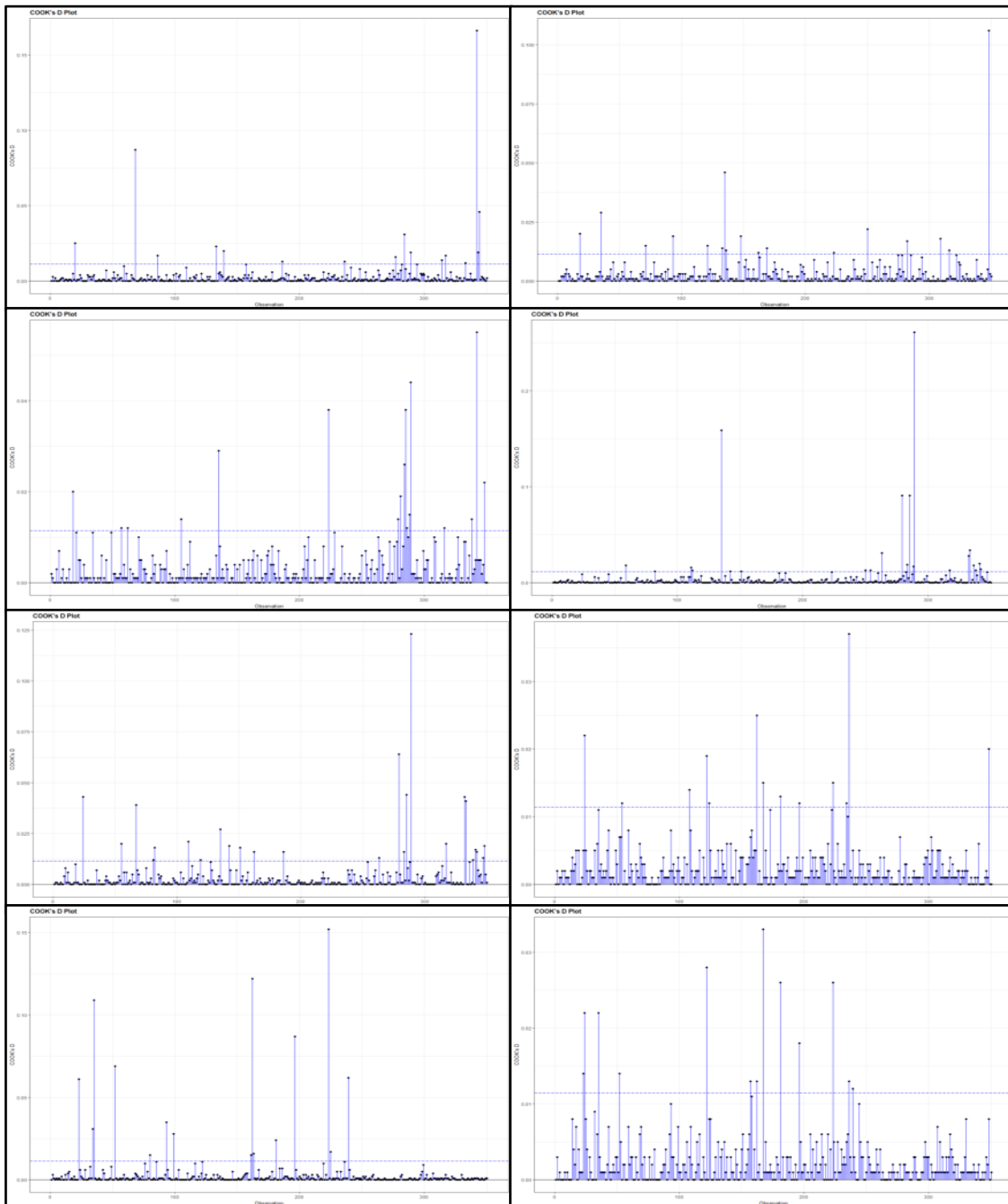


Figura 12. Distancias Cook para cada uno de los modelos estudiados. De izquierda a derecha y de arriba abajo: Palabras, Colores, Palabra-Color, Interferencia, Categorías, Perseveraciones, total Errores y SDMT.

#### 4.1.2 Efectos de las variables demográficas.

##### 4.1.2.1 Prueba de Stroop.

Los modelos de Regresión Múltiple fueron significativos para las puntuaciones de la prueba Stroop ( $p < 0.001$ ). Todos los modelos finales para las puntuaciones de la prueba Stroop se pueden encontrar en la Tabla 10.

Tabla 10.

Modelos de regresión final para las puntuaciones de la prueba de Stroop (N= 430)

Puntuaciones		B	Std. Error	$\beta$	t	Valor p	R <sup>2</sup> Ajustado	f <sup>2</sup>
Palabras	(Intercepto)	86.520	0.967		89.457	<0.001	0.282	0.393
	Edad	-0.421	0.051	-0.379	-8.210	<0.001		
	Educación	1.313	0.193	0.314	6.795	<0.001		
Colores	(Intercepto)	64.284	1.078		59.650	<0.001	0.335	0.504
	Edad	-0.414	0.041	-0.456	-9.976	<0.001		
	Edad <sup>2</sup>	-0.007	0.002	-0.154	-3.425	<0.001		
	Educación	1.056	0.152	0.309	6.955	<0.001		
PC	(Intercepto)	36.987	0.981		37.696	<0.001	0.366	0.577
	Edad	-0.309	0.037	-0.371	-8.310	<0.001		
	Educación	0.979	0.173	0.312	5.665	<0.001		
	Educación <sup>2</sup>	-0.063	0.022	-0.137	-2.812	<0.005		
	Sexo	0.794	1.592	0.022	0.499	<0.618		
	Educación X Sexo	9.8E-01	3.1E-01	0.179	3.188	<0.002		
Interferencia	(Intercepto)	0.523	0.569		0.920	<0.358	0.124	0.142
	Edad	-0.124	0.030	-0.210	-4.123	<0.001		
	Educación	0.567	0.114	0.255	4.992	<0.001		

El Modelo de Regresión final para la puntuación Stroop Palabras, presentó un efecto lineal en función de la edad, donde las puntuaciones disminuyeron a mayor número de años de edad (ver Figura 13). Además, se encontró que las puntuaciones Stroop Palabra incrementan linealmente en función de la educación. Las variables sexo e interacciones no afectaron esta puntuación. Los predictores del modelo explicaron el 28.2% del desempeño de los participantes en las puntuaciones de Stroop Palabras, según se observa en la Tabla 10.

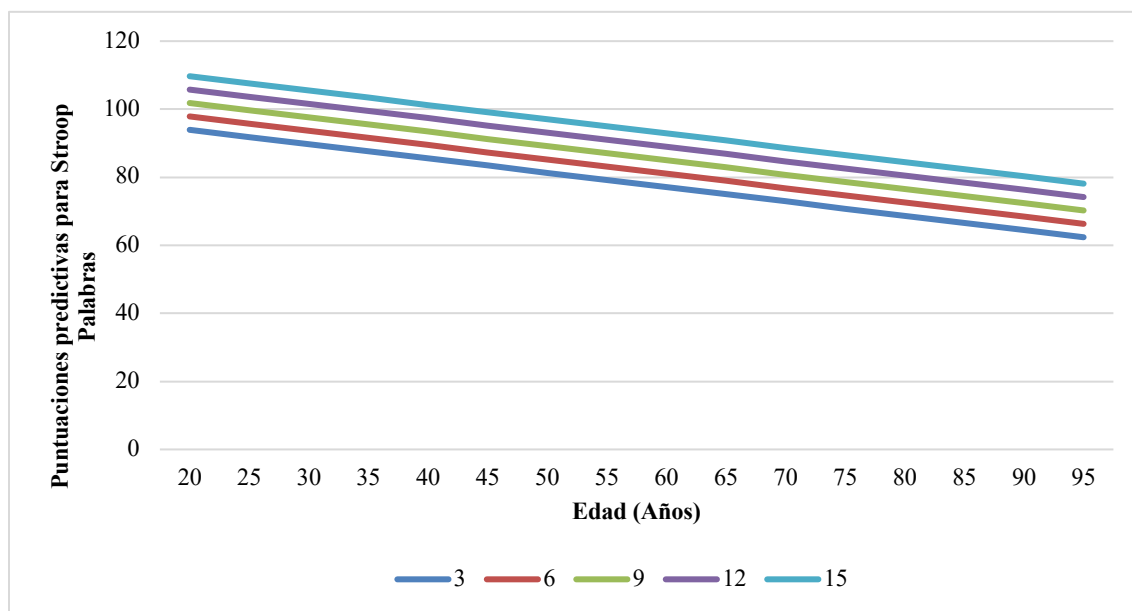


Figura 13. Gráficos del Modelo de Regresión final para la puntuación Stroop Palabras, en función de la edad. Las líneas de colores corresponden a años de escolaridad.

El Modelo de Regresión final para la puntuación Stroop Colores fue afectado por la edad cuadrática, donde las puntuaciones se mantienen estables hasta los 35 años de edad, y seguidamente decremantan de manera curvilínea (ver Figura 14). Se encontró que las puntuaciones Stroop Colores incrementan linealmente en función de la educación. La variable sexo y demás interacciones no afectaron esta puntuación. Los predictores del modelo

explicaron el 33.5% del desempeño de los participantes en las puntuaciones de Stroop Colores, de acuerdo a lo que se observa en la Tabla 10.

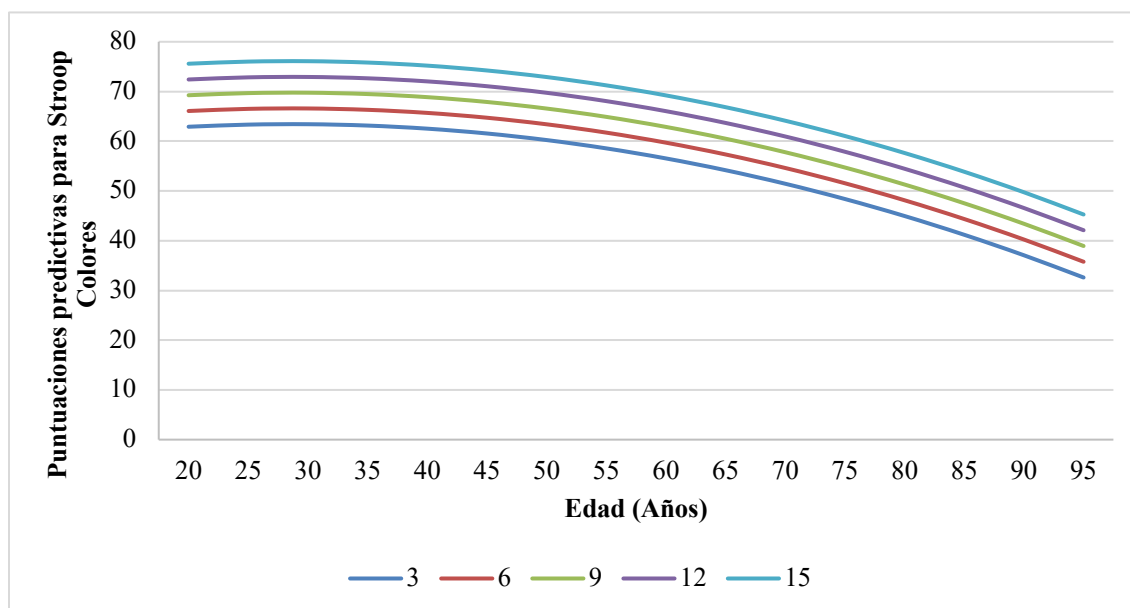


Figura 14. Modelo de Regresión final para la puntuación Stroop Colores por la edad cuadrática. Las líneas de colores corresponden a años de escolaridad.

El Modelo de Regresión final para la puntuación Stroop Palabra-Color fue afectado por la edad, donde las puntuaciones decremantan linealmente en función de la edad. Con respecto a la educación, se encontró que las puntuaciones Stroop Palabra-Color incrementan curvilíneamente en función de la educación. Finalmente, se encontró una interacción entre la educación cuadrática y el sexo, donde los hombres presentan menos puntuaciones que las mujeres hasta los nueve años de educación, seguidamente, las mujeres presentan menores puntuaciones que los hombres (ver Figura 15). Las otras interacciones no afectaron esta puntuación. Los predictores del modelo explicaron el 36.6% del desempeño de los participantes en las puntuaciones de Stroop Palabra-Color, de acuerdo a lo que se observa en la Tabla 10.



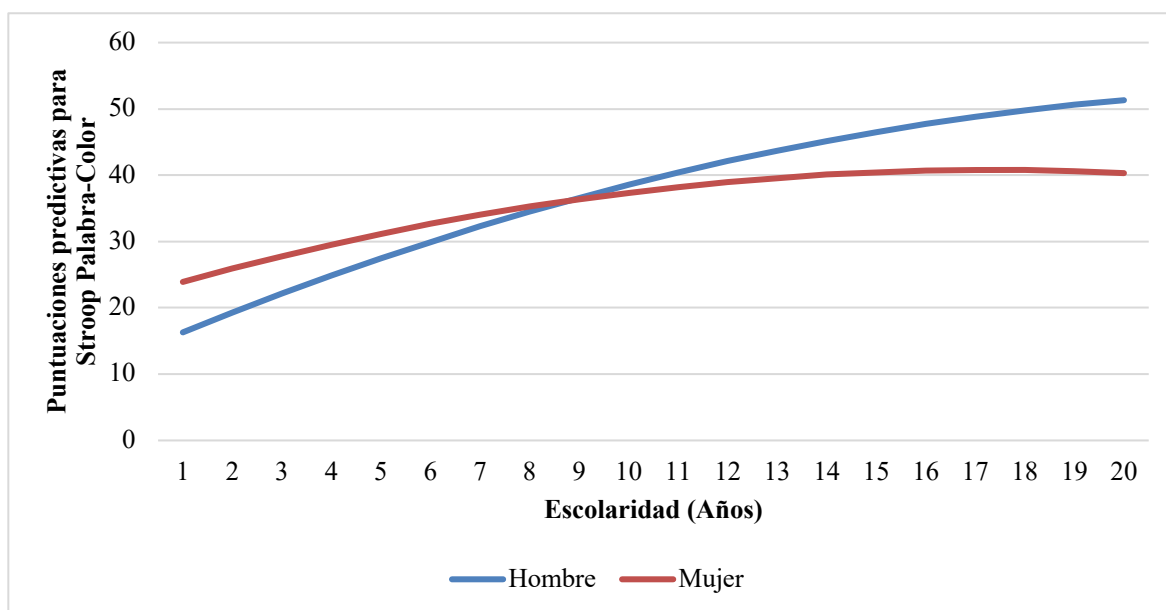


Figura 15. Puntuaciones predictivas comparando el rendimiento entre hombres y mujeres según educación en Stroop Palabra – Color.

El Modelo de Regresión final para la puntuación Stroop Interferencia presentó un efecto lineal en función de la edad, donde las puntuaciones decrecientan en función de los años de edad (ver Figura 16). Además, se encontró que las puntuaciones Stroop Interferencia incrementan linealmente en función de la educación. Las variables sexo e interacciones no afectaron esta puntuación. Los predictores del modelo explicaron el 12.4% del desempeño de los participantes en las puntuaciones de Stroop Interferencia, de acuerdo a lo que se observa en la Tabla 10.

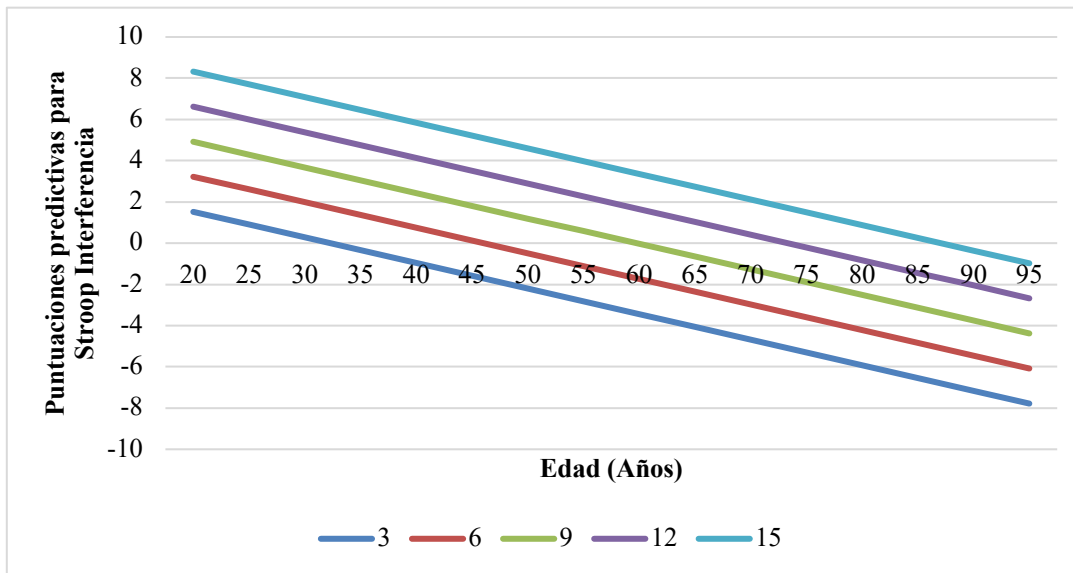


Figura 16. Modelo de Regresión final para la puntuación Stroop Interferencia en función de la edad. Las líneas de colores corresponden a años de escolaridad.

#### 4.1.2.2 M-WCST.

Los Modelos de Regresión Múltiple fueron significativos para las puntuaciones de la prueba M-WCST ( $p < 0.001$ ). Todos los modelos finales para las puntuaciones de la prueba M-WCST se muestran en la Tabla 11.

El Modelo de Regresión final para la puntuación Categorías presentó un efecto lineal en función de la edad, donde el número de categorías decremantan linealmente en función de los años de edad (ver Figura 17). Además, se encontró que el número de categorías incrementan linealmente en función de la educación. La variable sexo y demás interacciones no afectaron esta puntuación. Los predictores del modelo explicaron el 17.8% del desempeño de los participantes en las puntuaciones de Categorías, de acuerdo a lo que se observa en la Tabla 11.

Tabla 11.

Modelos de regresión final para las puntuaciones de M-WCST (N= 430)

Puntuaciones		B	Error estándar	$\beta$	t	Valor p	R <sup>2</sup> Ajustado	f <sup>2</sup>
(Intercepto)		4.338	0.091		47.508	<0.001		
Categorías	Edad	-0.024	0.005	-0.241	-4.875	<0.001	0.178	0.217
	Educación	0.115	0.018	0.311	6.301	<0.001		
(Intercepto)		4.117	0.497		8.280	<0.001		
Perseveración	Edad	0.145	0.019	0.372	7.552	<0.001	0.225	0.290
	Edad <sup>2</sup>	0.004	0.001	0.185	3.809	<0.001		
	Educación	-0.364	0.070	-0.249	-5.199	<0.001		
(Intercepto)		13.931	0.541		25.737	<0.001		
Total Errores	Edad	0.176	0.029	0.292	6.130	<0.001	0.234	0.305
	Educación	-0.772	0.108	-0.340	-7.139	<0.001		

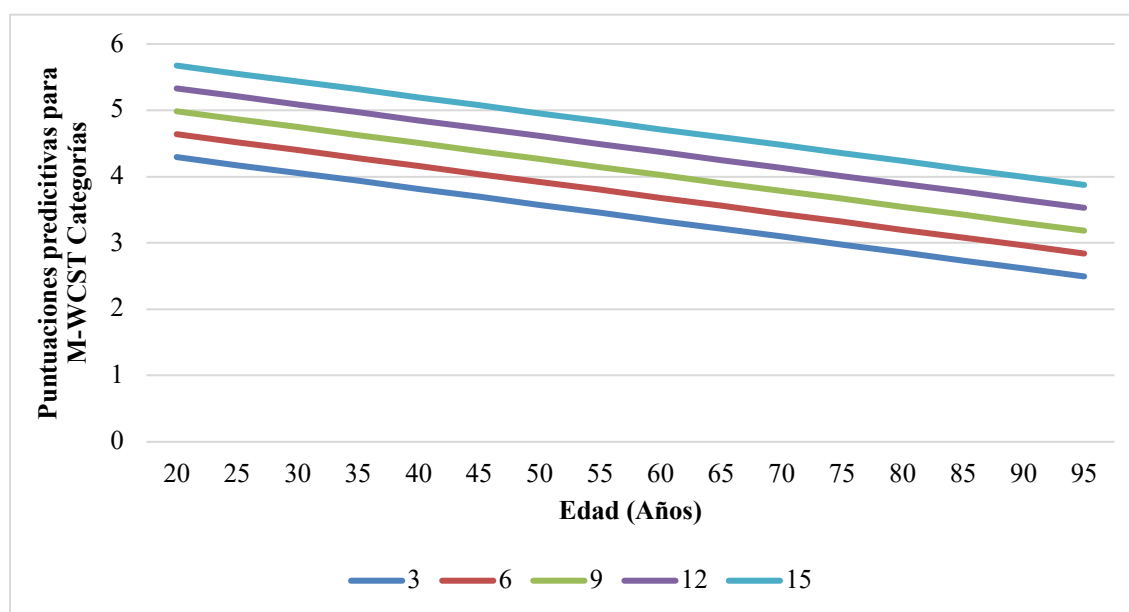


Figura 17. Modelo de Regresión final para la puntuación M-WCST Categorías en función de la edad. Las líneas de colores corresponden a años de escolaridad.

El Modelo de Regresión final para el número de Perseveraciones, muestra que las puntuaciones fueron afectadas por la edad cuadrática, donde el número de errores perseverativos decrecen hasta los 35 años de edad, y a partir de los 45 años de edad, el número de errores perseverativos incrementan de manera curvilínea en función de la edad (ver Figura 18). Se encontró que el número de errores perseverativos decrecientan linealmente en función de la educación. La variable sexo y demás interacciones no afectaron esta puntuación. Los predictores del modelo explicaron el 22.5% del desempeño de los participantes en las puntuaciones de errores perseverativos, de acuerdo a lo que se observa en la Tabla 11.

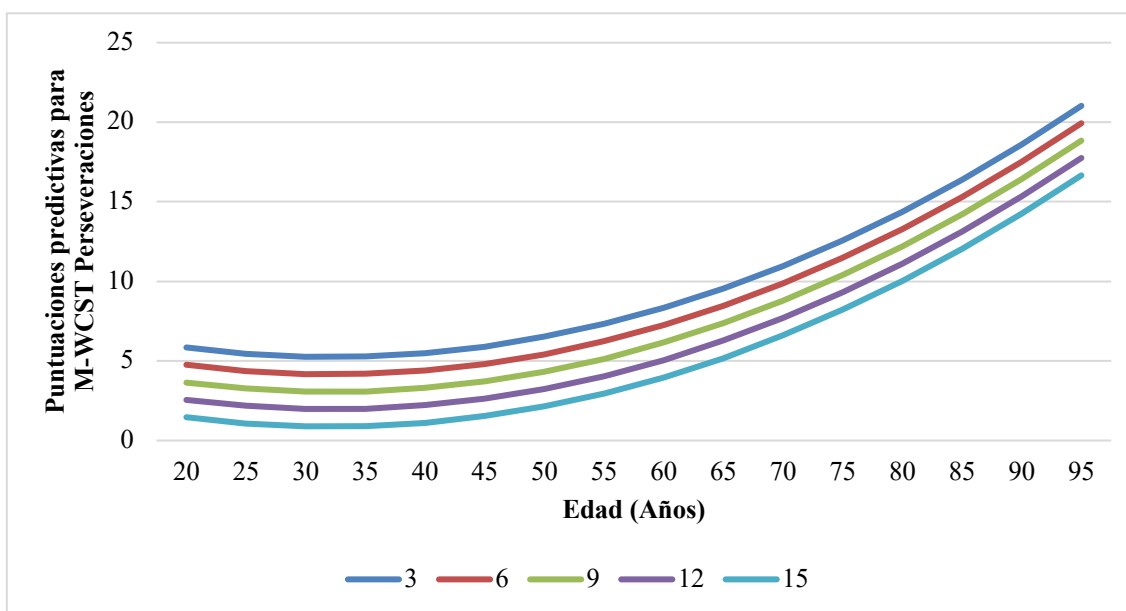


Figura 18. Modelo de Regresión final para el número de M-WCST perseveraciones por la edad cuadrática. Las líneas de colores corresponden a años de escolaridad.

El Modelo de Regresión final para el total de errores fue afectado por la edad, donde los errores se incrementan linealmente en función de la edad (ver Figura 19). Con respecto a la educación, se encontró que el número total de errores decrecientan linealmente en función

de la educación. Las variables sexo e interacciones no afectaron esta puntuación. Los predictores del modelo explicaron el 23.4% del desempeño de los participantes en las puntuaciones total de errores, de acuerdo a lo que se observa en la Tabla 11.

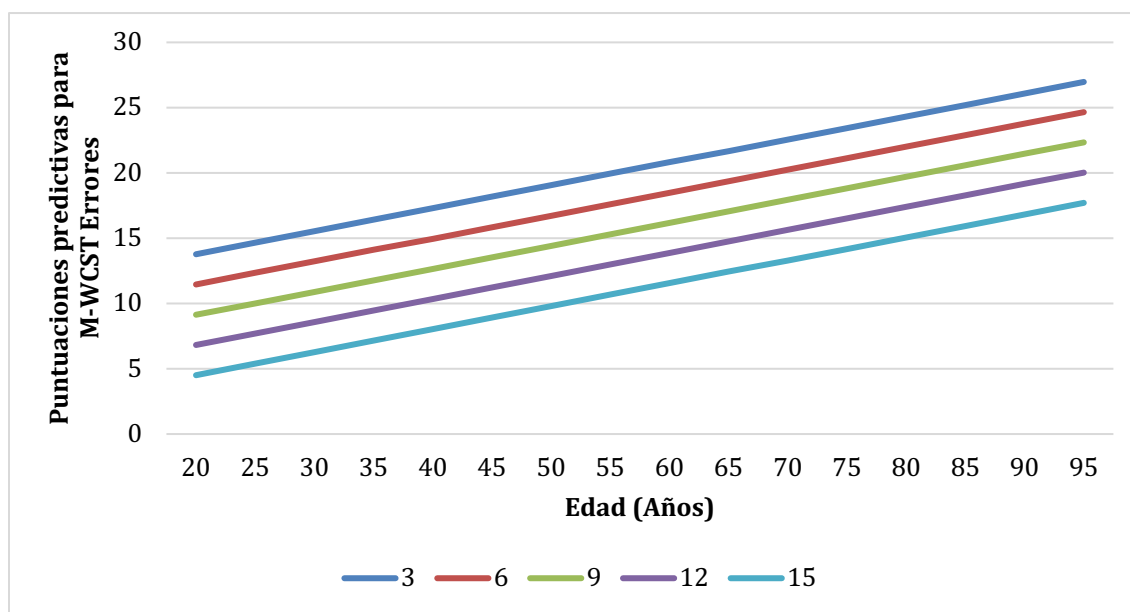


Figura 19. Modelo de Regresión final para M-WCST total de errores en función de la edad. Las líneas de colores corresponden a años de escolaridad.

#### 4.1.2.3 SDMT.

El modelo de Regresión Múltiple fue significativo para las puntuaciones de la prueba SDMT ( $p < 0.001$ ). El modelo final para las puntuaciones de la prueba SDMT se puede encontrar en la Tabla 12.

El Modelo de Regresión Múltiple final para la puntuación del SDMT muestra un efecto cuadrático de la edad, donde las puntuaciones decrecen curvilíneamente en función de la edad. Sin embargo, se encontró una interacción entre edad y la escolaridad, donde las puntuaciones decrecen de manera lineal cuando se tiene cinco o menos años de escolaridad. Mientras que las puntuaciones decrecen de manera curvilínea cuando se tiene de seis a

12 años de escolaridad (ver Figura 20). El modelo de regresión explica el 58.0% de la varianza de las puntuaciones SDMT, de acuerdo a lo que se observa en la Tabla 12.

Tabla 12.

Modelos de regresión final para el SDMT (N= 430)

Puntuación	B	Std. Error	$\beta$	t	Valor p	R <sup>2</sup> Ajustado	f <sup>2</sup>
(Intercepto)	40.835	0.999		40.869	<0.001		
Edad	-0.519	0.032	-0.621	-16.361	<0.001		
Edad <sup>2</sup>	-0.007	0.002	-0.166	-3.379	<0.001		
Educación	1.241	0.125	0.395	9.937	<0.001	0.580	1.381
Educación <sup>2</sup>	-0.110	0.023	-0.241	-4.900	<0.001		
Edad X Educación	-4.0E-04	6.7E-03	-0.002	-0.059	<0.953		
Edad <sup>2</sup> X Educación <sup>2</sup>	1.7E-04	5.8E-05	0.156	2.871	<0.004		

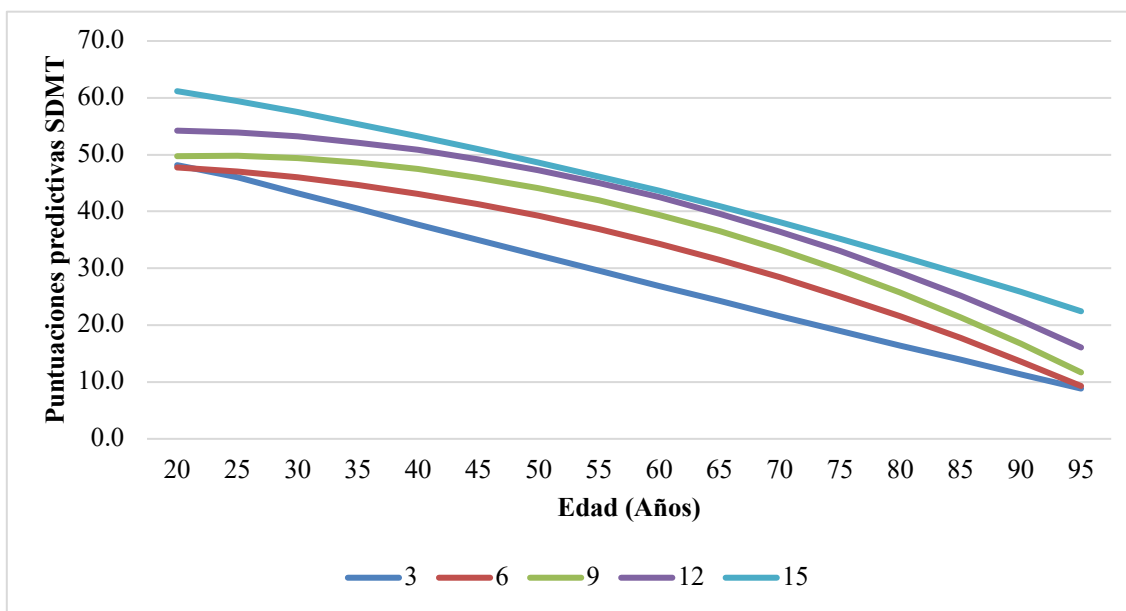


Figura 20. Modelo de Regresión Múltiple final para la puntuación del SDMT. Las líneas de colores corresponden a años de escolaridad.

#### 4.1.3 Datos normativos para las puntuaciones estudiadas.

Para el cálculo del percentil a partir de una puntuación directa, se puede seguir el siguiente ejemplo. Supóngase que se debe calcular el percentil para un hombre de 70 años de edad, con cinco años de escolaridad, que alcanzó una puntuación directa de 70 puntos en el Stroop Colores. Para obtener el percentil se deben seguir los siguientes cuatro pasos:

**Paso 1.** Usando la ecuación 1, se debe calcular el valor predictivo de la puntuación ( $\hat{Y}_i$ ) a partir de los parámetros ( $B$ ) establecidos en el modelo final de la tabla 10:

$$\hat{Y}_i = 64.284 + [-0.414 \cdot (70 - 50.3)] + [-0.007 \cdot (70 - 50.3)^2] + [1.056 \cdot (5 - 9.7)]$$

El resultado de la anterior ecuación da como valor predictivo  $\hat{Y}_i = 48.448$ .

**Paso 2.** Usando la ecuación 2, se debe calcular el valor residual ( $e_i$ ), calculando la diferencia entre la puntuación directa obtenida por la persona ( $Y_i$ ) y el valor predictivo calculado en el paso 1:

$$e_i = Y_i - \hat{Y}_i$$

$$e_i = 70 - 48.488$$

$$e_i = 21.552$$

**Paso 3.** El valor residual ( $e_i$ ) calculado anteriormente se debe estandarizar ( $z_i$ ). En este paso, se debe tener en cuenta que los valores residuales de los modelos finales para PC, Interferencia, Categorías, Perseveraciones y Errores no mostraron homocedasticidad, por tal motivo para calcular el valor de  $z_i$  de estas puntuaciones, se debe usar la  $DE_e$  de acuerdo al valor predictivo obtenido ( $\hat{Y}_i$ ) en el paso 1, tal como se muestra en la tabla 13. Una vez identificada el valor de la  $DE_e$  en la tabla 13, se debe usar la ecuación 3 para calcular  $z_i$ . En nuestro ejemplo, el valor de  $DE_e$  es igual a 14.164; despejando la fórmula:

$$z_i = e_i / DE_e$$

$$z_i = 21.552 / 14.164$$

$$z_i = 1.522$$

**Paso 4.** El valor estandarizado ( $z_i$ ) calculado en el paso 3 se puede transformar a su correspondiente percentil. Para ello, puede emplearse una calculadora en línea (p. ej., <http://www.measuringu.com/pcalcz/>) o una tabla de distribución acumulativa normal (p. ej., Strauss, Sherman, y Spreen, 2006). En cualquiera de los dos casos, se debe seleccionar el resultado que indique el porcentaje de área a una cola de la prueba. En este caso, el valor de  $z_i = 1.522$  corresponde a un percentil de 94.

En este último paso, se debe tener en cuenta que en los modelos finales para las puntuaciones del M-WCST (Categorías, Perseveraciones y Errores), sus valores residuales estandarizados no tuvieron una distribución normal, por tal motivo para transformar el valor estandarizado ( $z_i$ ) a percentil se debe usar la Tabla 14, donde  $z_i = -0.804$  en M-WCST categorías equivale a un percentil 25.



Tabla 13.

Desviaciones estándar de los valores residuales para cada uno de los modelos finales

<b>Puntuaciones</b>	<b>Valores predictivos (<math>\hat{Y}_i</math>)</b>	<b>(<math>DE_e</math>)</b>
Palabras	Todos los valores predictivos	18.041
Colores	Todos los valores predictivos	14.164
PC	$\leq 29.348$	9.434
	29.349 a 35.544	10.532
	35.545 a 42.858	11.306
	$\geq 42.859$	17.786
Interferencia	$\leq -2.631$	8.097
	-2.630 a 0.523	8.727
	0.524 a 3.703	9.995
	$\geq 3.704$	14.558
Categorías	$\leq 3.720$	1.839
	3.721 a 4.338	1.901
	4.339 a 4.987	1.769
	$\geq 4.988$	1.226
Perseveraciones	$\leq 2.416$	3.056
	2.417 a 5.024	4.532
	5.025 a 7.363	6.378
	$\geq 7.364$	10.065
Errores	$\leq 9.637$	6.985
	9.638 a 14.122	9.922
	14.123 a 18.276	11.414
	$\geq 18.277$	11.481
SDMT	Todos los valores predictivos	10.303

Tabla 14. Transformación de la puntuación  $z_i$  a percentil para Categorías, Perseveraciones y Errores

Percentil	Categorías	Perseveraciones	Errores	Percentil	Categorías	Perseveraciones	Errores
	$z_i$	$z_i$	$z_i$		$z_i$	$z_i$	$z_i$
<1	-3.076	-1.194	-1.713	50	0.294	-0.303	-0.205
1	-3.026	-1.144	-1.663	51	0.313	-0.292	-0.142
2	-2.154	-1.060	-1.393	52	0.337	-0.261	-0.116
3	-2.006	-1.024	-1.335	53	0.368	-0.234	-0.085
4	-1.908	-0.996	-1.307	54	0.412	-0.195	-0.062
5	-1.642	-0.971	-1.268	55	0.424	-0.189	-0.049
6	-1.593	-0.940	-1.229	56	0.443	-0.172	-0.040
7	-1.559	-0.932	-1.211	57	0.463	-0.164	-0.008
8	-1.521	-0.923	-1.186	58	0.478	-0.137	0.007
9	-1.463	-0.899	-1.153	59	0.518	-0.113	0.081
10	-1.450	-0.876	-1.133	60	0.518	-0.061	0.167
11	-1.407	-0.868	-1.117	61	0.537	-0.026	0.237
12	-1.341	-0.852	-1.102	62	0.553	0.024	0.276
13	-1.299	-0.829	-1.090	63	0.597	0.032	0.302
14	-1.191	-0.816	-1.061	64	0.604	0.050	0.329
15	-1.144	-0.805	-1.047	65	0.645	0.090	0.357
16	-1.115	-0.794	-1.022	66	0.665	0.120	0.397
17	-1.100	-0.791	-0.994	67	0.687	0.138	0.419
18	-1.079	-0.767	-0.984	68	0.691	0.175	0.443
19	-0.987	-0.762	-0.965	69	0.703	0.239	0.467
20	-0.966	-0.757	-0.946	70	0.748	0.268	0.505
21	-0.935	-0.741	-0.918	71	0.762	0.300	0.525
22	-0.901	-0.735	-0.889	72	0.778	0.315	0.554
23	-0.869	-0.717	-0.865	73	0.782	0.339	0.628
24	-0.849	-0.711	-0.841	74	0.801	0.401	0.671
25	-0.804	-0.689	-0.831	75	0.813	0.418	0.721
26	-0.725	-0.680	-0.815	76	0.821	0.458	0.762
27	-0.703	-0.672	-0.808	77	0.826	0.484	0.809
28	-0.646	-0.655	-0.776	78	0.840	0.515	0.883
29	-0.582	-0.643	-0.745	79	0.848	0.599	0.896
30	-0.511	-0.641	-0.725	80	0.866	0.666	0.935
31	-0.461	-0.632	-0.696	81	0.869	0.729	1.003
32	-0.424	-0.625	-0.674	82	0.880	0.774	1.083
33	-0.393	-0.615	-0.661	83	0.890	0.821	1.111
34	-0.356	-0.610	-0.640	84	0.901	0.859	1.131
35	-0.299	-0.603	-0.621	85	0.909	0.955	1.202
36	-0.292	-0.574	-0.595	86	0.926	1.092	1.278
37	-0.239	-0.537	-0.571	87	0.945	1.193	1.323
38	-0.195	-0.532	-0.544	88	0.966	1.234	1.356
39	-0.112	-0.516	-0.523	89	0.983	1.271	1.368
40	-0.085	-0.498	-0.492	90	0.999	1.341	1.409
41	-0.058	-0.482	-0.449	91	1.040	1.448	1.456
42	-0.033	-0.460	-0.412	92	1.076	1.515	1.484
43	0.045	-0.431	-0.385	93	1.113	1.613	1.661
44	0.134	-0.419	-0.353	94	1.136	1.715	1.758
45	0.148	-0.396	-0.327	95	1.170	1.852	1.875
46	0.178	-0.376	-0.308	96	1.282	2.471	2.001
47	0.232	-0.363	-0.301	97	1.349	2.776	2.180
48	0.248	-0.350	-0.269	98	1.387	3.128	2.266
49	0.278	-0.318	-0.263	99	1.487	3.465	2.682
				>99	1.822	5.000	4.218

#### 4.1.4 Calculadora de datos normativos.

El procedimiento de cuatro pasos para generar un dato normativo explicado anteriormente ofrece al clínico la capacidad de determinar un percentil exacto para un participante que tiene una puntuación específica en las puntuaciones estudiadas. Sin embargo, este método puede ser propenso a errores humanos debido al número de cálculos requeridos. Para mejorar la facilidad de uso, se elaboró una calculadora en Microsoft Excel en donde el clínico debe incluir la edad, la educación, sexo y las puntuaciones estudiadas para generar de manera automática el valor z y el percentil exacto (ver Figura 21). Esta herramienta está disponible como material suplementario en la página web del Instituto de Neuropsicología de Mexicali A.C. <https://www.inm-sinapsis.com/>

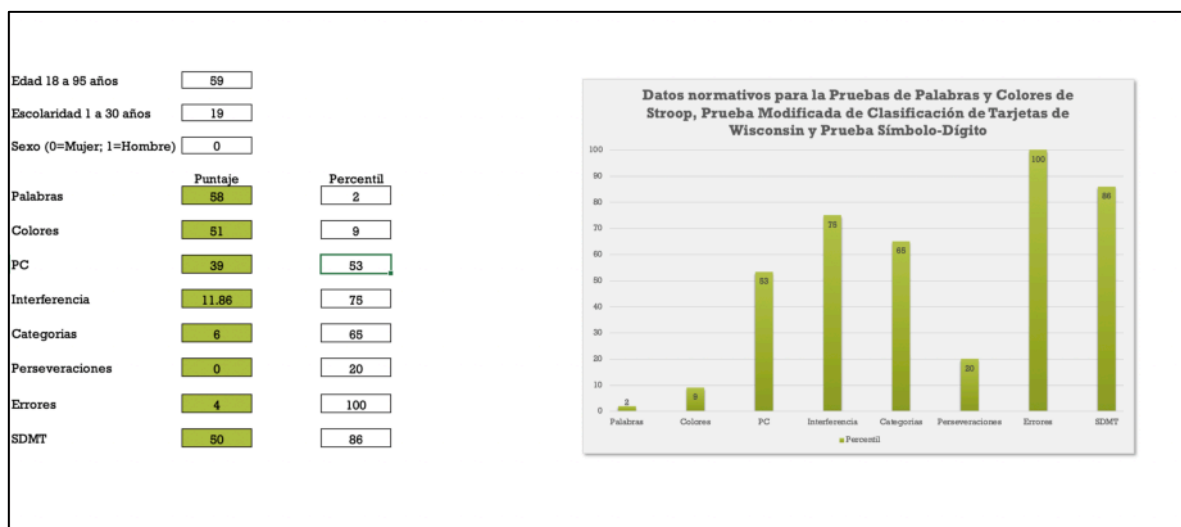


Figura 21. Calculadora de datos normativos para las tres pruebas utilizadas en este estudio. Se encuentran disponibles en <https://www.inm-sinapsis.com/>

#### 4.1.5 Comparaciones de modelos de regresión de 2015 y de 2020.

La submuestra 2 equivalente al 20% del total de la muestra de este trabajo ( $n=80$ ) se denominó Modelo 2020 y se comparó contra 80 sujetos de los estudios de datos normativos para la mismas pruebas publicados por Rivera et al.,(2015) y Arango-Lasprilla et al., (2015a y 2015b) denominados Modelo 2015. Hay que recordar que para este análisis las puntuaciones directas se transformaron a valores  $z_i$  en ambos modelos y luego se compararon con la prueba de los rangos signados de Wilcoxon.

De acuerdo con la Tabla 15 se encontraron diferencias significativas en todas las puntuaciones comparadas entre el modelo 2015 y el modelo 2020, excepto en las puntuaciones de Colores. Esto se puede observar en las figuras 22, 23 y 24 en las que se ven claramente áreas de traslape entre las puntuaciones de cada prueba, pero también se observan como es que existen zonas en las que sobresalen partes de las puntuaciones graficadas, específicamente en Categorías y Perseveraciones del M-WCST (Figura 23). En el caso de Errores del M-WCST y de SDMT (Figura 24) lo que se nota claramente es que casi no hay traslapes o superposiciones entre las puntuaciones de los modelos 2015 y 2020, especialmente en Errores del M-WCST.

Tabla 15.

Comparaciones de los valores  $z$  entre los modelos 2015 publicado y los modelos resultados de esta tesis

	Z	Sig. (2 colas)
Palabras 2015 - Palabras 2020	-2.83	0.005
Colores 2015 - Colores 2020	-0.671	0.502
PC 2015 - PC 2020	-2.173	0.030
Interferencia 2015 - Interferencia 2020	-4.753	<0.001
Categorías 2015 - Categorías 2020	-3.597	<0.001
Perseveraciones 2015 - Perseveraciones 2020	-7.77	<0.001
Errores 2015 - Errores 2020	-7.765	<0.001
SDMT 2015 - SDMT 2020	-2.83	<0.001

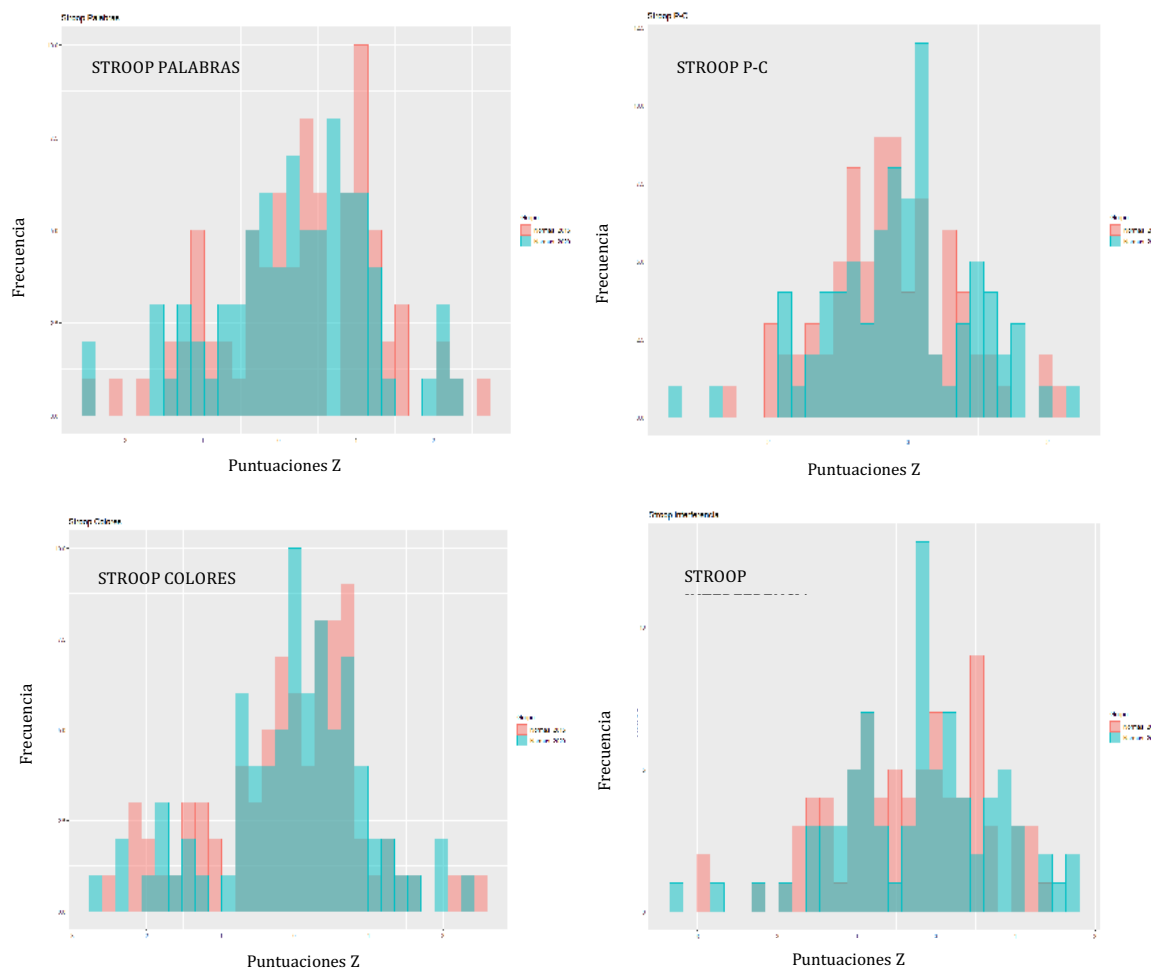


Figura 22. Comparación de los modelos 2015 y 2020 de las puntuaciones de la Prueba de STROOP.

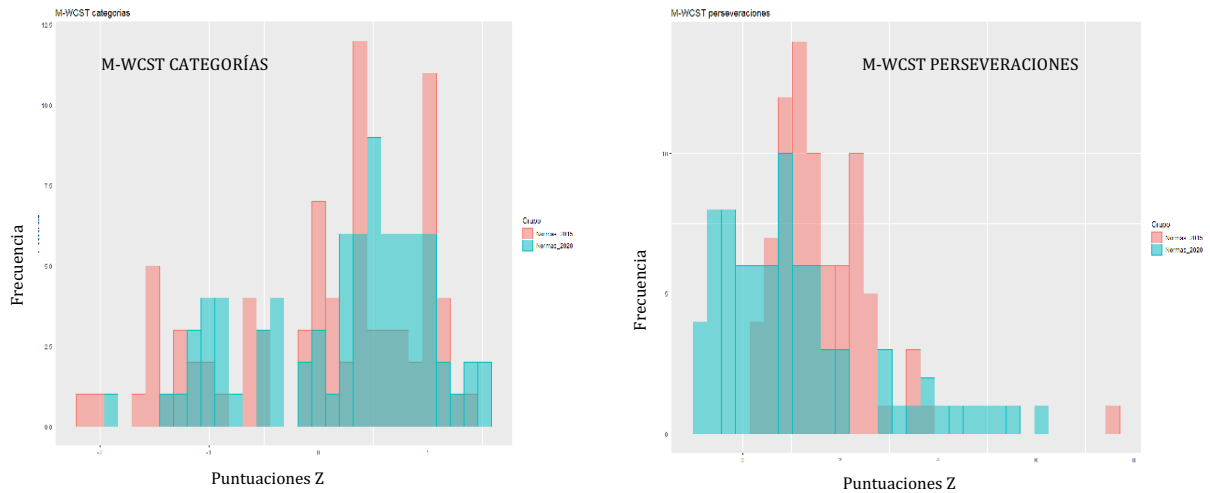


Figura 23. Comparación de los modelos 2015 y 2020 de las puntuaciones de Categorías y Perseveraciones de la Prueba M-WCST.

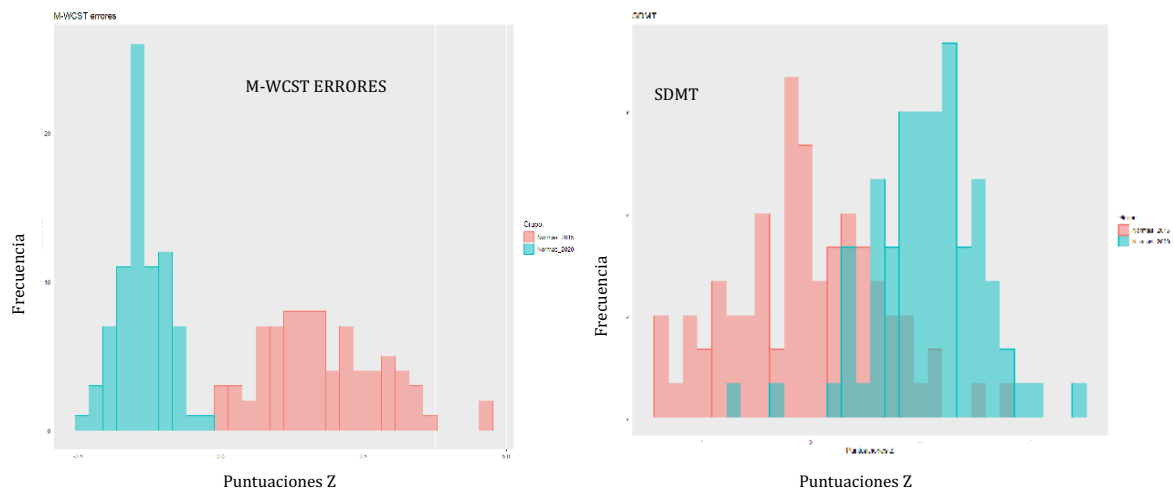


Figura 24. Comparación de los modelos 2015 y 2020 de las puntuaciones de Errores de la Prueba M-WCST y de la Prueba SDMT.

Al comparar los  $R^2$  de los modelos de regresión de 2015 y de 2020 lo que se encuentra es que en todas las pruebas aplicadas el modelo 2020 explica un porcentaje mayor de varianza que el modelo 2015. Por ejemplo, en la prueba en la que la diferencia entre  $R^2$  es menor es

la del total de Palabras de Stroop. La varianza explicada por el modelo 2015 es del 26.7%, en tanto que el modelo de 2020 explica el 28.6%, es decir una diferencia de 2 puntos porcentuales (Ver Tabla 16).

Tabla 16.

Retrospectiva comparando los  $R^2$  de los modelos de regresión de 2015 y 2020

<b>PUNTUACIÓN</b>		<b>2015 <math>R^2</math></b>	<b>2020 <math>R^2</math></b>
<b>STROOP</b>	Palabras	0.267	0.286
	Colores	0.306	0.340
	Palabra-Color	0.306	0.375
	Interferencia	0.072	0.129
<b>M-WCST</b>	Categorías	0.129	0.182
	Perseveración	0.125	0.238
	Errores	0.146	0.232
<b>SDMT</b>	Puntuación total	0.430	0.580

La prueba en la que se encontró una mayor diferencia entre las varianzas explicadas por cada modelo fue SDMT, ya que en 2015 la varianza explicada fue de 43% y en 2020 fue del 58%, es decir una diferencia de 15 puntos porcentuales.

## Capítulo 5. Discusión

### 5.1 Discusión

A manera de reflexión y de introducción a la discusión de esta tesis se presentarán ejemplos hipotéticos de puntajes obtenidos por personas de ciertas edades, géneros y escolaridades en los modelos 2015 y 2020 con el fin de comparar los percentiles obtenidos en cada modelo.

En la Figura 25 A se puede observar que para un puntaje de 27 en Stroop P-C un hombre de 20 años de edad y 1 de escolaridad obtiene, en el modelo 2020, un percentil de 55; en tanto que, para el mismo puntaje obtenido por un individuo con las mismas características, revisado con el modelo 2015 el percentil obtenido es de 5.

Si consideramos que esta misma persona logra un puntaje de 27 y tiene 20 años de edad, pero tiene 12 años de escolaridad, con el modelo 2020 obtendría un percentil de 3 y con el modelo de 2015 el percentil que obtendría es de 5 (Figura 25 A).

Finalmente, si el individuo tuviera 20 años de edad y 16 de escolaridad y lograra un puntaje de 27, con el modelo 2020 obtendría un percentil de 1, y con el modelo 2015 obtendría un percentil de 2 (Figura 25 A).



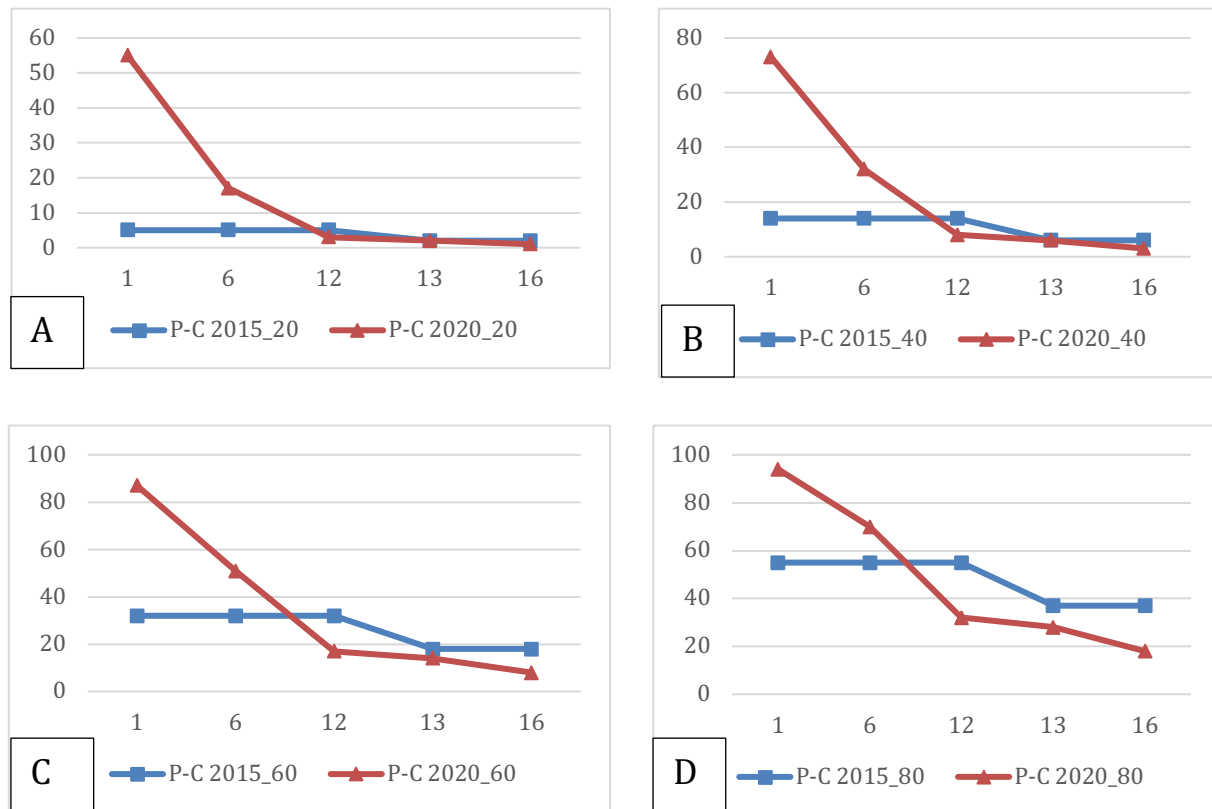


Figura 25. Comparación de los modelos de regresión 2015 vs 2015. Se muestran los percentiles obtenidos cuando un masculino obtiene un puntaje de 27 en STROOP P-C. A) 20 años de edad, B) 40 años de edad, C) 60 años de edad y D) 80 años de edad. En el eje de las Y se muestran los percentiles y en el eje de las X los años de escolaridad.

En el caso del número de errores totales en la prueba M-WCST el modelo final de regresión de 2020 contempla un intercepto de 13.931 con dos componentes adicionales que son Edad y Educación. Este modelo de regresión muestra una tendencia a obtener percentiles más bajos que el modelo de regresión 2015 en los niveles de educación de uno y seis años de escolaridad, en tanto que las escolaridades de doce, trece y dieciséis años en ambos modelos tienen valores muy similares.

Considerando que un hombre de 20 años de edad y 1 año de escolaridad cometiera 15 errores en total, de acuerdo con el modelo de regresión 2020 obtendría un percentil de 40,

pero si se revisara con el modelo de regresión del 2015 el percentil obtenido sería de 77. (Figura 26 A).

Suponiendo que una persona de 20 años de edad, pero con 12 años de escolaridad cometiera 15 errores en total y se revisara con el modelo de regresión de 2020, el percentil logrado sería de 86, pero si se revisara con el modelo de regresión de 2015, su percentil sería de 77 (Figura 26 A).

Si se revisara con el modelo de regresión 2020 a una persona de 20 años y 16 años de escolaridad que cometiera 15 errores, su percentil sería de 84, pero si se revisara con el modelo 2015 su percentil sería 87 (Figura 26 A).

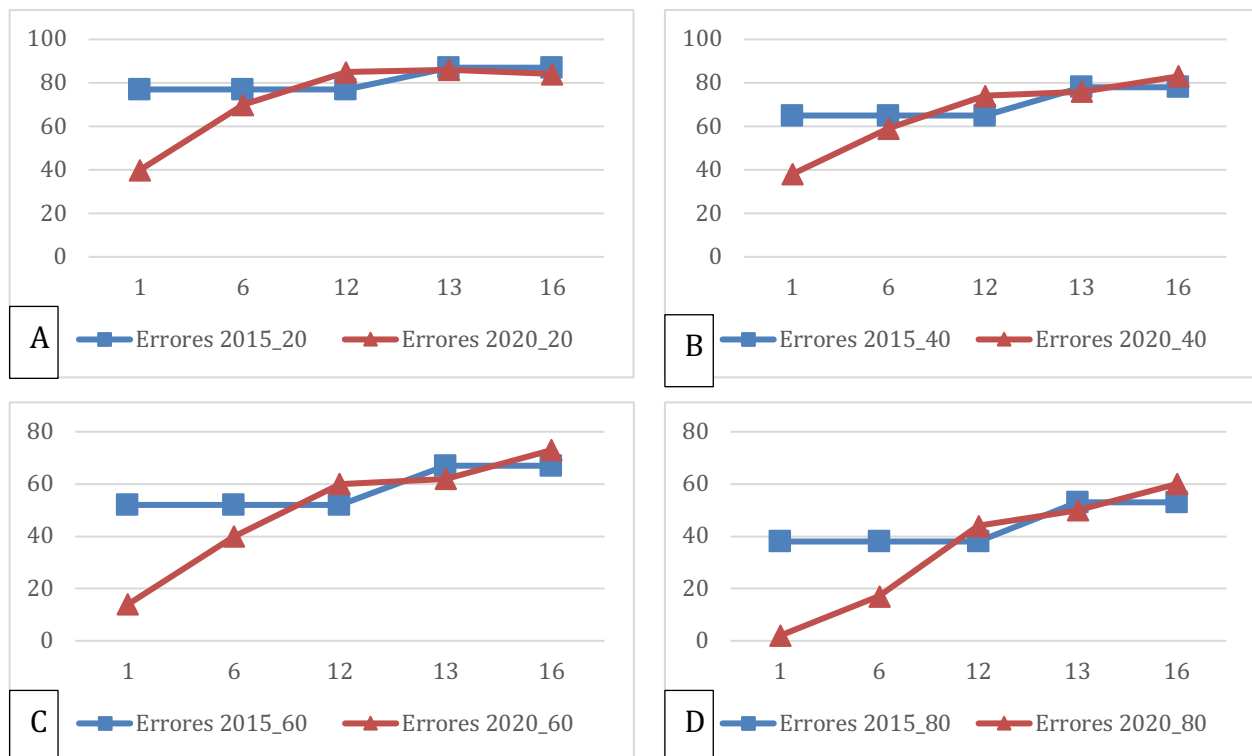


Figura 26. Comparación de los modelos de regresión 2015 vs 2020. Se muestran los percentiles obtenidos cuando un masculino comete 15 Errores en M-WCST. A) 20 años de edad, B) 40 años de edad, C) 60 años de edad y D) 80 años de edad. En el eje de las Y se muestran los percentiles y en el eje de las X los años de escolaridad.

En el caso del número de respuestas correctas en la prueba SDMT el modelo final de regresión de 2020 contempla un intercepto de 40.835 con seis componentes adicionales que son Edad,  $\text{Edad}^2$ , Educación,  $\text{Educación}^2$ ,  $\text{Edad} \times \text{Educación}$  y  $\text{Edad}^2 \times \text{Educación}^2$ . Este modelo de regresión muestra una tendencia a obtener percentiles más altos que el modelo de regresión 2015 en los niveles de educación de 1 y 6 años de escolaridad, en tanto que las escolaridades de 12, 13 y 16 años en ambos modelos tienen valores muy similares, aunque en ocasiones ligeramente más altos en el modelo de regresión de 2020, esto se observa en la Figura 27.

Considerando que un hombre de 20 años de edad y 1 año de escolaridad lograra 33 respuestas correctas, de acuerdo con el modelo de regresión 2020 obtendría un percentil de 18, pero si se revisara con el modelo de regresión del 2015 el percentil obtenido sería de 6. (Ver Figura 27 A).

Suponiendo que una persona de 20 años de edad, pero con 12 años de escolaridad obtuviera 33 respuestas correctas y se revisara con el modelo de regresión de 2020, el percentil logrado sería de 2, pero si se revisara con el modelo de regresión de 2015, su percentil sería de 6 (Ver Figura 27 A).

Si se revisara con el modelo de regresión 2020 a una persona de 20 años y 16 años de escolaridad que lograra 33 respuestas correctas, su percentil sería de 0, pero si se revisara con el modelo 2015 su percentil sería 2 (Ver Figura 27 A).

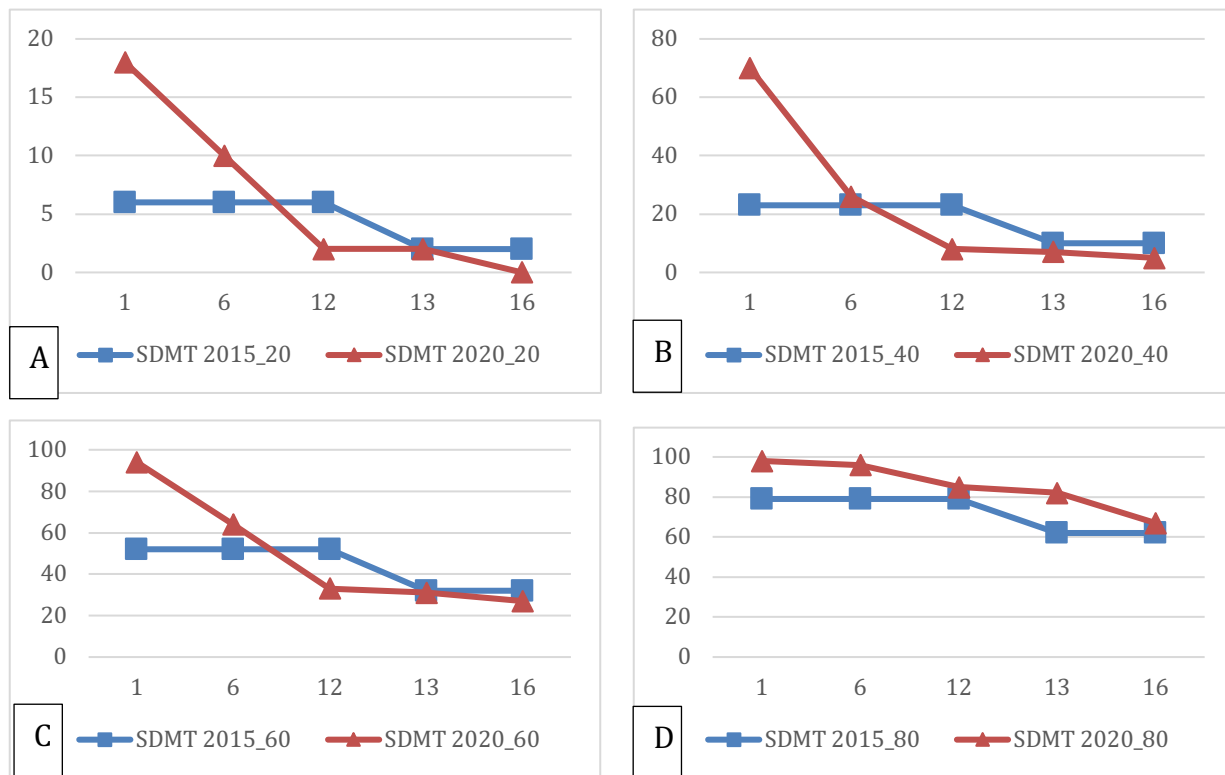


Figura 27. Comparación de los modelos de regresión 2015 vs 2020. Se muestran los percentiles obtenidos cuando un masculino obtiene 33 respuestas correctas en SDMT. A) 20 años de edad, B) 40 años de edad, C) 60 años de edad y D) 80 años de edad. En el eje de las Y se muestran los percentiles y en el eje de las X los años de escolaridad.

El objetivo de este trabajo fue obtener datos normativos para tres pruebas neuropsicológicas ampliamente utilizados en la clínica en población de Mexicali, Baja California. Los resultados muestran que los Modelos de Regresión Multivariada explican entre un 12.4% y un 36.6% de las puntuaciones en la Prueba de Stroop, entre el 17.8% al 23.4% de la varianza en las puntuaciones del Wisconsin y un 58.0% de la varianza de la Prueba de Símbolo Dígito.

Con respecto al Stroop se encontró que todas las puntuaciones (Palabras, Colores, Palabra-Color, Interferencia) de la prueba incrementan de forma lineal en función de la edad y de la educación. De forma que a mayor edad y mayor educación mejoraba la ejecución en todas las sub-escalas de la prueba. No se encontró ningún efecto del sexo con relación a ninguna de las puntuaciones excepto para la sub escala Palabra-Color donde se encontró una interacción entre la educación cuadrática y el sexo, donde los hombres presentan menos puntuaciones que las mujeres hasta los nueve años de educación, y posteriormente, las mujeres presentan menores puntuaciones que los hombres.

Los resultados del presente estudio concuerdan con otros estudios que han encontrado que la edad y la educación son variables que están íntimamente relacionadas con el desempeño de esta prueba (Rivera et al., 2015; Peña-Casanova et al., 2009; Vogel et al, 2013; Armengol, 2002; Oliveira, Mograbi, Gabrig, y Charchat-Fichman, 2016). Por ejemplo, en un estudio realizado por Llinas-Regla et al., (2013) con una muestra de 2151 adultos de entre 55 y 90 años de edad se encontró que las puntuaciones del Stroop mejoraban en función de la edad, es decir que los puntajes obtenidos por los grupos de 55, 60 y 65 años eran en promedio más altos que los obtenidos por los grupos de 75, 80 y 85 años de edad. En otro estudio realizado con niños de 7 a 12 años de edad (Roy, et al., 2018) se encontró que los niños de 7, 8 y 9 años de edad obtenían puntuaciones más bajas que los niños de 10, 11 y 12 años de edad. Por otra parte, los resultados de un estudio multicéntrico en el que se aplicó el Stroop a 5402 adultos de entre 18 y 90 años de edad en 12 países de Latinoamérica (Rivera et al., 2015) se encontró que la edad de la persona influye en la puntuación obtenida en la prueba de Stroop en todas las subpruebas con los grupos de menor edad obteniendo mejores puntuaciones que los grupos de mayor edad. La mejora fue un incremento lineal en la ejecución de la prueba de Stroop.

Con respecto al sexo los resultados del presente estudio coinciden con lo que han encontrado otros estudios en donde el sexo no es una variable que influye en el desempeño de esta prueba (Rivera et al., 2015). Sin embargo, se encontró una interacción entre la educación cuadrática y el sexo en el presente estudio para la variable palabra-color. Una posible explicación del por qué este hallazgo no ha sido reportado en otros estudios puede estar relacionado con que la variable educación cuadrática no ha sido incluida en los análisis que se realizaron (Rivera et al., 2015, Peña-Casanova et al., 2009; Vogel et al., 2013; Armengol, 2002; Oliveira, Mograbi, Gabrig, y Charchat-Fichman, 2016). Futuros estudios deberían incluir esta variable en el modelo de regresión para descartar o comprobar su influencia en las puntuaciones de la sub-escala de Palabra-Color.

Con respecto de la Prueba Modificada de Wisconsin se encontró que todas las puntuaciones (categorías, perseveraciones y número total de errores) de la prueba se incrementan de forma lineal en función de la edad y de la educación. De forma que a mayor edad, menor número de categorías, mayor número de respuestas perseverativas y mayor número de errores; mientras que a mayor educación, mayor número de categorías, menos errores perseverativos y menos cantidad total de errores. También se encontró que el número de perseveraciones fueron afectadas por la edad cuadrática, donde el número de errores perseverativos decrecen hasta los 35 años de edad, y a partir de los 45 años de edad, el número de errores perseverativos incrementan de manera curvilínea en función de la edad. No se encontró ningún efecto del sexo con relación a ninguna de las puntuaciones.

Los resultados del presente estudio concuerdan con otros estudios que han encontrado que la edad y la educación son variables que están íntimamente relacionadas con el desempeño de esta prueba (Cianchetti, Corona, Foscoliano, Contu y Sannio-Fancell, 2007; Fristoe, Salthouse y Woodward, 1997; Lin, Chan, Zheng, Yang, y Wang, 2007). Por ejemplo,

en un estudio realizado por Arango-Lasprilla et al. (2015) a adultos de 18 a 95 años de edad, en once países Latinoamericanos se encontró que, en general, las puntuaciones del Wisconsin mejoraban en función de la edad y de la escolaridad; es decir que a medida que incrementaba la edad y la escolaridad, mejoraban las puntuaciones y disminuían los errores.

Con respecto a la Prueba de Símbolo-Dígito se encontró que la cantidad de respuestas correctas de la prueba disminuyen en función de la edad y de la educación. De forma que a mayor edad y menor educación disminuye la ejecución en esta prueba (Peña-Casanova et al., 2009; Cancela, Ayán y Varela, 2012; Vanotti, et al., 2015; Burggraaff, Knol y Uitdehaag, 2017; Arango-Lasprilla et al., 2015b y 2017b).

Con respecto a la Prueba de Símbolo Dígito se encontró un efecto cuadrático de la edad, donde las puntuaciones decrecen curvilíneamente en función de la edad. Sin embargo, además se encontró una interacción entre edad y la escolaridad, donde las puntuaciones decrecen de manera lineal cuando se tiene cinco años o menos años de escolaridad. Mientras que las puntuaciones disminuyen de manera curvilínea cuando se tiene de seis a 12 años de escolaridad. No se encontró ningún efecto del sexo con relación a ninguna de las puntuaciones.

Al comparar los modelos de regresión generados en este estudio para 2015 y para 2020, lo que se encontró es que en el modelo 2015 no se consideraron las interacciones que pudieran existir entre todas las variables estudiadas, lo que resultó en un modelo con un menor poder para explicar la varianza de los datos. Solo se exploraron los efectos sobre la ejecución de tres variables: Edad, Escolaridad y Sexo, pero no el efecto cuadrático; ni las posibles interacciones entre estas. Otro aspecto relevante a revisar en el modelo de regresión de 2015 es que se asumió que los datos estaban distribuidos normalmente. Además de realizó

un corte en la escolaridad y los grupos se dividieron en dos: de 1 a 12 años de escolaridad y el otro con más de 12 años de escolaridad.

Todas estas consideraciones se tuvieron en cuenta en la elaboración de los modelos de regresión de 2020, de tal manera que algunas sub-pruebas, como en el caso de Palabra-Color en 2015 solo se conformó por el intercepto, la Edad y la Educación. En el modelo de regresión 2020 para esta misma variable, el modelo se conformó por el intercepto, Edad, Educación, Educación<sup>2</sup>, Sexo, y Educación x Sexo.

Adicionalmente, el modelo de regresión de 2010 no consideró la división de la muestra en grupos de edad, ni de escolaridad, de tal forma que los puntajes obtenidos por cada persona se comparaban contra todos los sujetos de la muestra, debido a que ya se consideraba el efecto que pudiera tener cada variable (Edad, Edad<sup>2</sup>, Educación, Educación<sup>2</sup>, Sexo, y todas las posibles interacciones entre estas variables). En los modelos de 2020 se asumieron los procedimientos sugeridos (Kutner, Natchtsheim, Neter y Li, 2015; Rivera et al, 2020) para evitar la multicolinealidad, la homocedasticidad, control de los Factores de Inflación de la Varianza. Además, se comprobó si los residuales estandarizados se distribuían normalmente para poder calcular y elaborar las tablas de percentiles.

Los modelos de regresión de 2020 son más confiables debido a que todas las consideraciones metodológicas y estadísticas que se menciona arriba, explican un mayor porcentaje de la varianza y los percentiles asignados a cada puntuación son más confiables.

## **5.2 Implicaciones**

Los resultados de este estudio tienen importantes implicaciones tanto a nivel clínico como investigativo. El hecho de contar con datos normativos para tres pruebas neuropsicológicas que evalúan los procesos de atención, funciones ejecutivas y velocidad de



procesamiento en personas de Mexicali, es algo muy importante ya que los clínicos podrán utilizar estas pruebas en su práctica habitual para evaluar estos procesos en una gran variedad de personas con alteraciones neuropsicológicas y psiquiátricas. En la actualidad existen pocos estudios en la literatura que se hayan realizado sobre baremación de pruebas neuropsicológicas en población mexicana. Los pocos estudios que existen en su gran mayoría han sido estudios con muestras muy pequeñas que han generado baremos basados en medias y desviaciones estándar, y en los cuales no se ha evaluado el papel que juegan variables tales como la edad, el sexo y la escolaridad, que son variables muy importantes a tener en cuenta a la hora de utilizar estas pruebas.

Este estudio también es el primer paso hacia la elaboración de futuras investigaciones que se puedan realizar con el objetivo de baremar muchas más pruebas neuropsicológicas que se utilizan en la región, y de las cuales hasta el día de hoy no contamos con baremos locales.

Al ser pruebas de uso común, fácil de administrar y de corta duración, el tener baremos de estas pruebas va a facilitar mucho más la práctica clínica y el trabajo del neuropsicólogo, puesto que dispondrá de tres herramientas muy útiles a la hora de poder realizar evaluaciones neuropsicológicas en esta población.

En la actualidad muchas enfermedades psiquiátricas y neurológicas conllevan la aparición de problemas de atención, dificultades de velocidad de procesamiento de la información y alteraciones en las funciones ejecutivas. El tener baremos de estas pruebas en personas sanas nos permitirá a futuro poder recoger información en una gran cantidad de poblaciones clínicas con el objetivo de crear puntos de corte y determinar la sensibilidad y la especificidad de estas pruebas en poblaciones clínicas.

### 5.3 Limitaciones

Los resultados de este estudio deben ser interpretados a la luz de las siguientes limitaciones: 1) Debido al alto número de inmigrantes con los que en la actualidad cuenta la ciudad es de señalar que las normas aquí presentadas no pueden ser utilizadas en personas con características socioculturales diferentes a la población con las que se llevó a cabo este estudio. 2) Debido a que todas las personas que participaron vivían en Mexicali, los resultados de este estudio no se pueden generalizar a personas que vivan en otras ciudades de Baja California. Futuros estudios deberían llevarse a cabo para poder generar datos normativos de estas tres pruebas en otras ciudades de Baja California con el objetivo de obtener una muestra más representativa. 3) En este estudio se controlaron variables como la edad, años de escolaridad y el sexo; sin embargo, otras variables que pueden estar en relación con el desempeño cognitivo tales como el nivel de bilingüismo, el nivel educativo de los padres o la raza no se incluyeron como parte de los análisis puesto que no se recogió dicha información. 4) Este estudio se realizó con personas que tenían por lo menos un año de escolaridad, por tal motivo los resultados de este estudio no deberían utilizarse en población analfabeta. 5) Las personas que participaron en este estudio fueron personas que no tenían antecedentes de enfermedades neurológicas, psiquiátricas o historia de problemas de aprendizaje o consumo de sustancias. Sin embargo, esta información fue obtenida a través del auto reporte de cada uno de los participantes. Futuros estudios deberían utilizar medidas más objetivas con el propósito de poder descartar que estas personas realmente no hayan tenido ninguno de estos problemas que se han mencionado anteriormente. 6) La gran mayoría de personas que participaron en este estudio provenía de zonas urbanas de la ciudad, futuros estudios deberían incluir también población de zonas rurales. 7) En futuros estudios se deben de incluir muestras clínicas y estimar puntos de corte y otros aspectos psicométricos

complementarios a este estudio. 8) De acuerdo a los análisis de rachas las puntuaciones del M-WCST no se distribuyen aleatoriamente, sin embargo, los análisis de los supuestos estadísticos y su respectivo control, en la regresión se pueden utilizar para el control de sesgos, y 9) El muestreo usado en este estudio fue no probabilístico, lo cual impide la estimación del grado de generalización de los resultados a toda la población adulta de Mexicali. Por otra parte, el principal problema de muestreos no probabilístico es que la precisión es indeterminada, sin embargo, el tamaño de la muestra total y la presencia de unidades muestrales en los diferentes rangos de edad, escolaridad y sexo, la fiabilidad y la evidencia de validez de los test, así como la similitud de los resultados con otros estudios de datos normativos, hace que este estudio tenga validez interna y sus datos puedan ser usados por neuropsicólogos.

## Capítulo 6. Conclusión

El objetivo de este trabajo de tesis fue obtener datos normativos para tres pruebas neuropsicológicas (Prueba de Colores y Palabras de Stroop, Prueba Modificada de Clasificación de Tarjetas de Wisconsin y Prueba de Modalidades Símbolo-Dígito) ampliamente utilizados en la clínica en población de Mexicali, Baja California. Como conclusión se puede decir que;

- 1) En todas las pruebas y sub-pruebas se demostró el efecto de la edad y de la escolaridad.

Aunque solo en la sub-pruebas de Palabra-Color se encontró un efecto en la interacción del sexo combinado con la educación cuadrática. Entonces las variables que tienen un efecto mayor en la ejecución de todas las pruebas aplicadas son la edad y la educación. A menor edad generalmente se encontraron mayor cantidad de aciertos y menor cantidad de errores, en tanto que con la educación ocurrió lo contrario, y las personas con mayor escolaridad obtenían menos errores y mayor cantidad de aciertos.

- 2) El uso de la técnica de regresión múltiple produce datos normativos más confiables, si se controlan los efectos de factores como la colinealidad, homocedasticidad, multicolinealidad, factor de inflación de la varianza, ya que al tener esto en cuenta se puede trabajar con muestras más pequeñas, se evita dividir la muestra en grupos de edad, de escolaridad o de sexo; de tal manera que la discrepancia entre los estadísticos de la muestra y los parámetros de la población real se reduce y también disminuye la probabilidad de errores en la toma de decisiones clínicas.

- 3) Se logró obtener los datos normativos en la Prueba de Colores y Palabras de Stroop, Prueba Modificada de Clasificación de Tarjetas de Wisconsin y Prueba de Modalidades Símbolo-Dígito para la población de Mexicali, Baja California. Esto contribuirá a que

los neuropsicólogos de esta entidad puedan contar con parámetros normativos generados en el lugar de origen y de vivienda de sus pacientes.

- 4) Es necesario continuar con la obtención de datos normativos para otras pruebas neuropsicológicas que evalúen otros procesos cognoscitivos y también realizar este tipo de estudios en poblaciones con características clínicas como Alzheimer, Alteraciones Atencionales con o sin Hiperactividad, Síndrome Metabólico, Trauma de Cráneo, por mencionar algunas.

## Referencias

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- Amato, M. P., Portaccio, E., Goretti, B., Zipoli, V., Ricchiuti, L., De Caro, M. F... Trojano, M. (2006). The Rao's Brief Repeatable Battery and Stroop Test: normative values with age, education and gender corrections in an Italian population. *Multiple Sclerosis Journal*, 12(6), 787-793.
- Anstey, K. J., Matters, B., Brown, A. K., y Lord, S. R. (2000). Normative data on neuropsychological tests for very old adults living in retirement villages and hostels. *The clinical neuropsychologist*, 14(3), 309-317.
- Arango-Lasprilla, J. C., y Rivera, D. (2015a). *Neuropsicología en Colombia. Datos normativos, estado actual y retos a futuro*. Colombia: Universidad Autónoma de Manizales.
- Arango-Lasprilla, J. C., Olabarrieta, L., Rivera, D., Olivera, S., De los Reyes, C., Calderón, A... Utria, O. (2015b). *Situación actual de la neuropsicología en Colombia*. En: Arango-Lasprilla, J. y Rivera, D., (Eds) *Neuropsicología en Colombia: Datos normativos, estado actual y retos a futuro*. Universidad Autónoma de Manizales, Colombia.
- Arango-Lasprilla, J. C., Rivera, D., Longoni, M., Saracho, P., Garza, M., Aliaga, A... Perrin, P. (2015c). Modified Wisconsin Card Sorting Test (M-WCST): Normative data for the Latin American Spanish speaking adult population *NeuroRehabilitation* 37, 563–590.

- Arango-Lasprilla, J. C., Rivera, D., Rodríguez, G., Garza, M., Galarza-del-Angel, J., Rodríguez, W.... Perrin, P. (2015d). Symbol Digit Modalities Test: Normative data for the Latin American Spanish speaking adult population *NeuroRehabilitation* 37, 625–638.
- Arango-Lasprilla, J. C., Rivera, D., Trapp, S., Jiménez-Pérez, C., Hernández, C., Pohlenz, F... Albaladejo-Blázquez, N. (2017a). Symbol Digit Modalities Test: Normative data for Spanish-speaking pediatric population. *NeuroRehabilitation*, 41, 639–647
- Arango-Lasprilla, J. C., Stevens, L., Morlett-Paredes, A., Ardila, A., y Rivera, D. (2017b). Profession of neuropsychology in Latin America *Appl Neuropsychol* 24(4): 318-330.
- Ardila, A., Pineda, D., y Rosselli, M. (2000). Correlation Between Intelligence Test Scores and Executive Function Measures. *Archives of Clinical Neuropsychology*, 15(1) 31-36.
- Arffa, S., Lovell, M., Podell, K., y Goldberg, E. (1998). Wisconsin Card Sorting Test Performance in Above Average and Superior School Children: Relationship to Intelligence and Age. *Archives of Clinical Neuropsychology*, 13(8) 713–720.
- Armengol, C. G. (2002). Stroop test in Spanish: Children's norms. *The Clinical Neuropsychologist*, 16 (1), 67-80.
- Banich, M., Milham, M., Atchley, R., Cohen, N., Webb, A., Wszalek, T... Magin, R. (2000). FMRI Studies of Stroop Tasks Reveal Unique Roles of Anterior and Posterior Brain Systems in Attentional Selection. *Journal of Cognitive Neuroscience* 12(6) 988–1000.
- Beato, R., Amaral-Carvalho, V., Guimaraes, C., Tumas, V., Souza, P., Oliveira, G., y Caramelli, P. (2012). Frontal assessment battery in a Brazilian sample of healthy controls: Normative data. *Arquivos de Neuro-Psiquiatria*, 70(4), 278-280.

- Beltrán, C., y Solís, G. (2012). Evaluación Neuropsicológica en Adolescentes: Normas para Población de Bucaramanga. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias*, 12(2), 77-93.
- Ben-David, B., Nguyen, L., y van Lieshout1, P. (2011). Stroop Effects in Persons with Traumatic Brain Injury: Selective Attention, Speed of Processing, or Color-Naming? A Meta-analysis. *Journal of the International Neuropsychological Society*, 17, 354–363.
- Benedict, R., Mezhir, J., Walsh, K., y Hewitt, R. (2000). Impact of Human Immunodeficiency Virus Type-1-Associated Cognitive Dysfunction on Activities of Daily Living and Quality of Life. *Archives of Clinical Neuropsychology*, 15(6) 535-544.
- Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Annual Review of Psychology*, 47(1), 563-592.
- Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, 39(1), 15-22.
- Bezdicek, O., Lukavsky, J., Stepankova, H., Nikolai, T., Axelrod, B., Michalec, J... Miloslav Kopecek, M. (2015). The Prague Stroop Test: Normative standards in older Czech adults and discriminative validity for mild cognitive impairment in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology*, 37(8), 794–807.
- Brugnolo, A., De Carli, F., Accardo, J., Amore, M., Bosia, L., Bruzzaniti, C... Girtler, N. (2016). An updated Italian normative dataset for the Stroop color word test (SCWT). *Neurological Sciences* 37(3), 365–372.



- Burggraaff, J., Knol, D., y Uitdehaag, B. (2017). Regression-Based Norms for the Symbol Digit Modalities Test in the Dutch Population: Improving Detection of Cognitive Impairment in Multiple Sclerosis? *European Neurology* 77(5-6), 246–252.
- Butman, J., Allegri, R. F., Harris, P., y Drake, M. (2000). Fluencia verbal en español. Datos normativos en Argentina. *Medicina*, 60(5/1), 561-564.
- Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., y Venneri, A. (2004). Modified card sorting test: Normative data. *Journal of Clinical and Experimental Neuropsychology*, 26(2), 246-250.
- Cancela, J., Ayán, C., y Varela, S. (2012). Valores normativos del “Symbol Digit Modalities Test” de aplicación en poblaciones españolas residentes en geriátricos: un estudio piloto *Actas Esp Psiquiatr*, 40(6), 299-303.
- Carley, L., Kydd, R., Robertson, M., Pillai, A., McNair, N., Lee, N... Russell, B. (2015). Acute effects of the designer drugs benzyloperazine (BZP) and trifluoromethylphenylpiperazine (TFMPP) using functional magnetic resonance imaging (fMRI) and the Stroop task. A pilot study. *Psychopharmacology*, 232(16), 2969–2980 .
- Cassiani-Miranda, C., Pérez-Anibal, E., Vargas-Hernández, M., Herazo-Bustos, M., y Cabarcas\_Tovar, A. (2018). Validez de apariencia y adaptación de la escala PHQ-9 para la detección de sintomatología depresiva en universitarios de ciencias de la salud de Cartagena (Colombia). *Salud Uninorte, Barranquilla (Col)*, 34 (1), 75-87.
- Chaytor, N., Barbosa-Leiker, C., Ryan, C., Germine, L., Hirsh, I., y Weinstock, R. (2019). Clinically significant cognitive impairment in older adults with type 1 diabetes. *Journal of Diabetes and Its Complications*, 33(1), 91–97.

- Christman, A., Vahorsdall, T., Pearlson, G., Hill-Briggs, F., y Schretlen, D. (2010). Cranial Volume, Mild Cognitive Deficits, and Functional Limitations Associated with Diabetes in a Community Sample. *Archives of Clinical Neuropsychology* 25(1), 49–59.
- Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., y Sannio-Fancello, G. (2007). Modified Wisconsin Card Sorting Test (M-WCST, MWCST): Normative data in children 4-13 years old, according to classical and new types of scoring. *The Clinical Neuropsychologist*, 21(3), 456-578.
- Comité de Planeación para el Desarrollo del Estado de Baja California, COPLADE. (2017). *Publicaciones sociodemográficas, Baja California, Mexicali*. Recuperado de <http://www.copladebc.gob.mx/publicaciones/2017/Mensual/Mexicali%202017.pdf>
- Damakis, G. (2003). A Meta-Analytic Review of the Sensitivity of the Wisconsin Card Sorting Test to Frontal and Lateralized Frontal Brain Damage. *Neuropsychology*, 17(2), 255–264.
- Diez-Quevedo, C., Rangil, T., Sánchez-Planell, L., Kroenke, K., y Spitzer, R. (2001). Validation and utility of the Patient Health Questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. *Psychosomatic Medicine*, 63, 679-686.
- Diforio, D., Walker, E. F., y Kestler, L. P. (2000). Executive functions in adolescents with schizotypal personality disorder. *Schizophrenia Research*, 42(2), 125-134.
- Drake, A., Weinstock-Guttman, B., Morrow, S., Hojnacki, D., Munschauer, F., y Benedict, R. (2010). Psychometrics and normative data for the Multiple Sclerosis Functional

- Composite: replacing the PASAT with the Symbol Digit Modalities Test. *Multiple Sclerosis* 16(2) 228–237.
- Duchek, J., Balota, D., Thomas, J., Snyder, A., Rich, P., Benzinger, T... Ances, B. (2013). Relationship Between Stroop Performance and Resting State Functional Connectivity in Cognitively Normal Older Adults. *Neuropsychology*, 27(5), 516–528.
- Dymowski, A., Owens, J., Pansford, J., y Willmott, C. (2015). Speed of processing and strategic control of attention after traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 37(10), 1024–1035.
- Estevez-Gonzalez, A., Garcia-Sanchez, C. y Barraquer-Bordas, Ll. (2000) Los lóbulos frontales: el cerebro ejecutivo. *Revista de Neurología*, 31 (6): 566-577.
- Fernandes Lopes, R., y De Lima Agrimon, I. (2010). Idosos com diabetes mellitus tipo 2 e o desempenho cognitivo no teste Wisconsin de classificação de cartas (WCST). *Universitas Psychologica*, 9(3), 697-713.
- Folstein, M., Folstein, S., McHugh, P. R., y Fanjiang, G. (1975). Examen Cognoscitivo Mini-Mental. *Trial: Professional Manual. Psychological Assessment Resources*.
- Fontes de Gracia, S., García-Gallego, C., Quintanilla Cobián, L., Rodríguez Fernández, R., Rubio de Lemus, P., & Sarriá Sánchez, E. (2018). Fundamentos de investigación en psicología. Madrid: Editorial Universidad Nacional de Educación a Distancia.
- Fristoe, N., Salthouse, T., y Woodward, J. (1997). Examination of Age-Related Deficits on the Wisconsin Card Sorting Test. *Neuropsychology*, 11(3), 428-436
- Golden, C. J. (2010). *Stroop, Test de Colores y Palabras: Manual (5a ed.)*. Madrid: Tea Ediciones.
- Goldstein, L., y McNeil, J. (2004). Clinical neuropsychology: a practical guide to assessment and management for clinicians. Wiley and Sons, England.

- González-Blanch, C., Medrano, L., Muñoz-Navarro, R., Ruiz-Rodríguez, P., Moriana, J., Limonero, J...Cano-Vindel, A. (2018). Factor structure and measurement invariance across various demographic groups and over time for the PHQ-9 in primary patients in Spain. *PLoS ONE* 13 (2): e0193356. <https://doi.org/10.1371/journal.pone.0193356>
- Gonzalez, H., Whitfield, K., West, B., Williams, D., Lichtemberg, P., y Jackson, J. (2007). Modified-Symbol Digit Modalities Test for African Americans, Caribbean Black Americans, and non-Latino Whites: Nationally representative normative data from the National Survey of American Life. *Archives of Clinical Neuropsychology*, 22(5), 605–613.
- Grant, D. A., y Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology*, 38(4), 404-411.
- Greve, K., Bianchini, K., Hartley, S., y Adams, D. (1999). The Wisconsin Card Sorting Test in Stroke Rehabilitation: Factor Structure and Relationship to Outcome. *Archives of Clinical Neuropsychology*, 14(6), 497–509
- Guardia-Olmos, J., Blanxart, M. F., Pero-Cebollero, M., & Turbany-Oset, J. (2007). Analisis de datos en psicología (1.a ed., pp. 100–102). Delta publicaciones.
- Guardia-Olmos, J., Però-Cebollero, M., Rivera, D., y Arango-Lasprilla, J. (2015). Methodology for the development of normative data for ten Spanish-language neuropsychological tests in eleven Latin American countries. *NeuroRehabilitation*, 37, 493-499
- Guisse, B., Thompson, M., Greve, K., Bianchini, K., y West, L. (2014). Assessment of performance validity in the Stroop Color and Word Test in mild traumatic brain injury

- patients: A criterion-groups validation design. *Journal of Neuropsychology*, 8(1), 20–33
- Hammar, S., Sørensen, L., Ardal, G., Oedegaard, K., Kroken, R., Roness, A., y Lund, A. (2010). Enduring cognitive dysfunction in unipolar major depression: A test–retest study using the Stroop paradigm Scandinavian. *Journal of Psychology*, 51, 304–308
- Hammes, J. (1973). *De Stroop Kleur-Woord Test: Handleiding [The Stroop Color-Word Test: Manual]*. Amsterdam: Swets y Zeitlinger.
- Hancock, P., y Larner, A. (2009). Clinical utility of Patient Health Questionnaire-9 (PHQ-9) in memory clinics. *International Journal of Psychiatry in Clinical Practice*, 13 (3), 188-191
- Hebben, N., y Milberg, W. (2011). *Fundamentos para la evaluación neuropsicológica*. Manual Moderno, México.
- Heflin, L., Laluz, V., Jang, J., Ketelle, R., Miller, B., y Kramer, J. (2011). Let’s Inhibit Our Excitement: The Relationships Between Stroop, Behavioral Disinhibition, and the Frontal Lobes. *Neuropsychology*, 25(5), 655–665
- Helland, T., y Asbjørnsen, A. (2000). Executive functions in dyslexia. *Child Neuropsychology*, 6(1), 37-48
- Hernández-Sampieri, R., Fernández, C., y Baptista, P. (2003). *Metodología de la investigación, Tercera Edición*. México:Mc Graw Hill.
- Herd, S., Banich, M., y O’Reilly, R. (2006). Neural Mechanisms of Cognitive Control: An Integrative Model of Stroop Task Performance and FMRI Data. *Journal of Cognitive Neuroscience*, 18(1), 22–32

- Hubbard, E., y Motl, R. (2015). Sedentary behavior is associated with disability status and walking performance, but not cognitive function, in multiple sclerosis. *Appl. Physiol. Nutr. Metab.*, 40, 203–206
- Igarashi, K., Oguni, H., Osawa, M., Awaya, Y., Kato, M., Mimura, M., y Kashima, H. (2002). Wisconsin card sorting test in children with temporal lobe epilepsy. *Brain and Development*, 24(3), 174-178.
- International Test Commission (2014). *El uso de los tests y otros instrumentosa de evaluación en la investigación*. España: International Test Commission.
- Instituto Nacional de Estadística Geografía e Informática (2016) *Panorama sociodemográfico de Baja California 2015*. México. Recuperado de [http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod\\_serv/contenidos/espanol/bvinegi/productos/nueva\\_estruc/inter\\_censal/panorama/702825082093.pdf](http://internet.contenidos.inegi.org.mx/contenidos/Productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/inter_censal/panorama/702825082093.pdf)
- Jodzio, K., y Biechowska, D. (2010). Wisconsin Card Sorting Test as a Measure of Executive Function Impairments in Stroke Patients. *Applied Neuropsychology*, 17, 267–277
- Kapoula, Z., Lê, T., Bonnet, A., Bourtoire, P., Demule, E., Fauvel, C... Yang, Q. (2010). Poor Stroop performances in 15-year-old dyslexic teenagers. *Exp Brain Res*, 203, 419–425.
- Kohrt, B., Luitel, N., Acharya, P., y Jordans, M. (2016), Detection of depression in low resource settings: validation of the Patient Health Questionnaire (PHQ-9) and cultural concepts of distress in Nepal. *BMC Psychiatry*, 16:58 <https://doi.org/10.1186/s12888-016-0768-y>

- Kho, C., Lu, W., Chen, H., Hsueh, P., Hsieh, J., y Hsieh, C. (2011). Test–Retest Reliability and Practice Effect of the Oral-format Symbol Digit Modalities Test in Patients with Stroke. *Archives of Clinical Neuropsychology*, 26, 356–363
- Kiely, K., Butterworth, P., Watson, N., y Woodens, M. (2014). The Symbol Digit Modalities Test: Normative Data from a Large Nationally Representative Sample of Australians. *Archives of Clinical Neuropsychology*, 29, 767 – 775
- Kroenke, K., Spitzer, R., y Williams, J. (2001). The PHQ-9. *Journal of Gen Internal Medicine*, 16 (9), 606-613.
- Kuccu, D., Kandemir, M., Unal, A., Topcular, B., y Kirbas, D. (2012). Longitudinal Study of Cognitive Impairment in Multiple Sclerosis: A 5-year Follow-up. *Archives of Neuropsychiatry*, 49:29- 32.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models (5th ed.)*. New York: McGraw Hill.
- Landry, y Al-Taie, (2016). A Meta-analysis of the Wisconsin Card Sort Task in Autism. *J Autism Dev Disord* 46:1220–1235
- Lee, S., Kim, D., Sohn, M., Lee, J., Lee, S., Shin, Y... Kim, Y. (2020). Determining the cut-off score for the Modified Barthel Index and the Modified Rankin Scale for assessment of functional independence and residual disability after stroke. *PLoS ONE* 15 (1): e0226324.<https://doi.org/10.1371/journal.pone.0226324>
- Lee, H. C., Falkmer, T., y Chee, D. Y. (2011). Performance of drivers with parkinson’s disease under the effect of cognitive overloading: insinuation for assessment and training. *Accid Anal Prevent*.
- Lezak, M. (1995). *Neuropsychological Assessment* (3th ed.). New York: Oxford University Press.

- Lezak, M., Howieson, D., y Loring, D. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.
- Lin, H., Chan, R., Zheng, L., Yang, T., y Wang, Y. (2007). Executive functioning in healthy elderly Chinese people. *Archives of Clinical Neuropsychology*, 22, 501–511
- Llinas-Regla, J., Vilalta-Franch, J., López-Pousa, S., Calvo-Perxas, L., y Garre-Olmo, J. (2013). Demographically Adjusted Norms for Catalan Older Adults on the Stroop Color and Word Test. *Archives of Clinical Neuropsychology*, 28, 282–296
- Luo, C., Proctor, R., y Weng, X. (2015). A Stroop effect emerges in the processing of complex Chinese characters that contain a color/related radical. *Psychological Research*, 79, 221–229
- Magalhaes, S. S., y Hamdan, A. C. (2010). The Rey Auditory Verbal Learning Test: Normative data for the Brazilian population and analysis of the influence of demographic variables. *Psychology y Neuroscience*, 3(1), 85.
- Mahoney, F., y Barthel, D. (1965). Functional evaluation: the Barthel Index. *Maryland State Medical Journal*, 14, 56-61.
- Martins, I. P., Castro-Caldas, A., Townes, B. D., Ferreira, G., Rodrigues, P., Marques, S., . . . & Derouen, T. (2005). Age and sex differences in neurobehavioral performance: A study of Portuguese elementary school children. *International Journal of Neuroscience*, 115(12), 1687-1709.
- Mcleod, C. (1991). Canadian Psychology. *Psychologie canadienne*, 32(3), 521-524.
- Melrose, R., Young, S., Weissberger, G., Natta, L., Harwood, D., Mark Mandelkern, M., y Sultzer, D. (2017). Cerebral metabolic correlates of attention networks in Alzheimer's Disease: A study of the Stroop. *Neuropsychologia*, 106, 383–389



- Moering, R. G., Schinka, J. A., Mortimer, J. A., y Graves, A. B. (2004). Normative data for elderly African Americans for the Stroop color and word test. *Archives of Clinical Neuropsychology*, 19(1), 61-71.
- Moran, L., & Yeates, K. O. (2011). Stroop Color and Word Test, Children's Version. In *Encyclopedia of Clinical Neuropsychology* (pp. 2403-2404). Springer New York.
- Ospina Botero, D. (2001). *Introducción al muestreo*. Editorial Universidad Nacional De Colombia.
- Morgan, A. B., y Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review*, 20(1), 113-136.
- Miyake, A., Friedman, N., Emerson, M., Witzki, A., y Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49-100
- Negahban, H., Ebrahimzade, M., y Mehravar, M. (2017). The effects of cognitive versus motor demands on postural performance and weight bearing asymmetry in patients with stroke *Neuroscience Letters* 659, 75-79
- Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex*, 12(4), 313-324.
- Obonsawin, M. C., Crawford, J. R., Page, J. Chalmers, P., Low, G., y Marsch, P. (1999). Performance on the Modified Card Sorting Test by normal, healthy individuals: Relationship to general intellectual ability and demographic variables. *Journal of Clinical Psychology*, 38(1), 27-41.
- Oliveira, R. M., Mograbi, D. C., Gabrig, I. A., y Charchat-Fichman, H. (2016). Normative data and evidence of validity for the Rey Auditory Verbal Learning Test, Verbal

- Fluency Test, and Stroop Test with Brazilian children. *Psychology & Neuroscience*, 9(1), 54.
- Olivera, S., Rivera, D., De los Reyes, C., Quijano, M., Calderón, J., Utria, O... Arango-Lasprilla, J. C. (2015). Datos normativos del test de Colores y Palabras (Stroop) para población Colombiana. En: Arango-Lasprilla, J.C. y Rivera, D. (Eds). *Neuropsicología en Colombia. Datos normativos, estado actual y retos a futuro*. Colombia: Universidad Autónoma de Manizales.
- Ostrosky-Solís, F., Ardila, A., y Rosselli, M. (1999). NEUROPSI: A brief neuropsychological test battery in Spanish with norms by age and educational level. *Journal of the International Neuropsychological Society*, 5(05), 413-433.
- Ostrosky-Solís, F., Dávila, G., Ortiz, X., Vega, F., Ramos, G. G., de Celis, M... Molina, B. (1999). Determination of normative criteria and validation of the SKT for use in Spanish-speaking populations. *International Psychogeriatrics*, 11(2), 171-180
- Ozonoff, S. (1995). Reliability and Validity of the Wisconsin Card Sorting Test in Studies of Autism. *Neuropsychology*, 9(4), 491-500
- Padilla, L., Téllez, A., Galarza, J., Téllez, H., Garza, M. y Garza, C. (2016) *Diccionario de neuropsicología*. México: El Manual Moderno.
- Papazian, O., Alfonso, I. y Luzondo, R. (2006) Trastornos de las funciones ejecutivas. *Revista de Neurología*, 42 (Supl 3): s45-s50.
- Park, M., y Lee, S. (2019). Effect of a dual-task program with different cognitive tasks applied to stroke patients: A pilot randomized controlled trial. *NeuroRehabilitation*, 44, 239–249.
- Peña-Casanova, J., Quiñones-Úbeda, S., Gramunt-Fombuena, N., Quintana, M., Aguilar, M., Molinuevo, J. L... y Antúnez, C. (2009a). Spanish Multicenter Normative Studies

- (NEURONORMA Project): norms for the Stroop color-word interference test and the Tower of London-Drexel. *Archives of Clinical Neuropsychology*, 24(4), 413-429.
- Peña-Casanova, J., Quiñones-Ubede, S., Quintana-Aparicio, M., Aguilar, M., Badenes, D., Molinuevo, J... Blesa, R. (2009b). Spanish Multicenter Normative Studies (NEURONORMA Project): Norms for Verbal Span, Visuospatial Span, Letter and Number Sequencing, Trail Making Test, and Symbol Digit Modalities Test *Archives of Clinical Neuropsychology* 24, 321–341
- Perlstein, W., Carter, C., Barch, D., y Baird, J. (1998). The Stroop task and attention deficits in schizophrenia: A critical evaluation of card and single-trial Stroop methodologies *Neuropsychology*, 12, (3) 414 -425
- Pineda, D. (2000) La función ejecutiva y sus trastornos. *Revista de Neurología*, 30 (8): 764-768.
- Portellano, J. (2005). Introducción a la neuropsicología. Mc Graw-Hill: Interamericana, España.
- Proulx, M., y Elmasry, H. (2015). Stroop interference in adults with dyslexia. *Neurocase*, 21(4) 413–417
- Raina, S., Chander, V., Raina, S. y Grover, A. (2016). *Post-hoc* principal component analysis on a largely illiterate elderly population from North-west India to identify important elements of mini-mental state examination. *Journal of Neuroscience in Rural Practice*, 7, 44-47
- Regard, M. (1981). *Cognitive rigidity and flexibility: A neuropsychological study*. Unpublished Ph.D. dissertation, University of Victoria.

- Rivera, D., Morlett-Paredes, A., Peñalver A., Irías M., Soto-Añari, M., Aguayo, A...  
Arango-Lasprilla, J. C. (2017). Stroop Color-Word Interference Test: Normative data for Spanish-speaking pediatric population. *NeuroRehabilitation*, 41 605–616
- Rivera, D., Olabarrieta-Landa, L., Van der Elst, W., Gonzalez, I., Ferre-Cascales, R., Peñalver, A., Rodriguez-Lorezana, A., Galarza-del-Angel, J., Irias, M., y Arango-Lasprilla, J. C. (2020). Regression-based normative data for children from Latin America: Phonological verbal fluency letters M, R, and P. Assessment, En Prensa
- Rivera, D., Perrin, P., Stevens, L., Garza, M., Weil, C., Saracho, C., Rodriguez, W., Rodriguez-Agudelo, Y., Rábago, B., Weiler, G., García de la Cadena, C., Longoni, M., Martínez, C., Ocampo-Barba, N., Aliaga, A., Galarza-del-Angel, J. Guerra, A., Esenarro, L. y Arango-Lasprilla, J. C. (2015) Stroop Color-Word Interference Test: Normative data for the Latin American spanish speaking adult population. *NeuroRehabilitation*, 37 : 591-624.
- Rivera, D., Olabarrieta-Landa, L., Van der Elst, W., Gonzalez, I., Rodríguez-Agudelo, Y., Aguayo-Arelis, A., ... & Arango-Lasprilla, J.C. (2019). Normative data for verbal fluency in healthy Latin American adults: letter M, and fruits and occupations categories. *Neuropsychology*, 33(3), 287-300. doi: 10.1037/neu0000518.
- Rivera, D., Olabarrieta-Landa, Van der Elst, W., Gonzalez, I., Ferrer Cascales, R., Peñalver Guia, A.I. ... & Arango-Lasprilla, J.C. (2020). Regression-Based Normative Data for Children from Latin America: Phonological Verbal Fluency Letters M, R, and P. Assessment. doi:10.1177/1073191119897122
- Rockers, K., Ousley, O., Sutton, T., Schoenberg, E., Coleman, K., Walker, E., y Cubells, J. F. (2009). Performance on the Modified Card Sorting Test and its relation to

- psychopathology in adolescents and young adults with 22q11.2 deletion syndrome. *Journal of Intellectual Disability Research*, 53(7), 665-676.
- Rodríguez, L., Pulido, N., y Pineda, C. (2016). Propiedades psicométricas del Stroop, test de colores y palabras en población colombiana no patológica. *Universitas Psychologica*, 15(2), 255-272.
- Rosselli, M., Ardila, A., Florez, A., y Castro, C. (1990). Normative data on the Boston Diagnostic Aphasia Examination in a Spanish-speaking population. *Journal of Clinical and Experimental Neuropsychology*, 12(2), 313-322.
- Roy, A., Kefi, M., Bellaj, T., Fournet, N., Le Gall, D., y Roulin, J. (2018). The Stroop test: A developmental study in a French children sample aged 7 to 12 years. *Psychologie Française* 63, 129–143
- Rybakowski, J., Borkowska, A., Czerski, P., Kapelski, P., Dimitrza-Weglarz, M., y Hauser, J. (2005). An association study of dopamine receptors polymorphisms and the Wisconsin Card Sorting Test in schizophrenia. *J Neural Transm*, 112: 1575–1582
- Sachs, M., Kaplan, J., Der Sarkissian, A., y Habibi, A. (2017). Increased engagement of the cognitive control network associated with music training in children during an FMRI Stroop task. *PLoS ONE* 12 (10): e0187254. <https://doi.org/10.1371/journal.pone.0187254>
- Sarno, M., Gaztanaga, W., Banerjee, N., Bure-Reyes, A., Rooks, J., Margolesky, J... Levin, B. (2019). Revisiting eligibility for deep brain stimulation: Do preoperative mood symptoms predict outcomes in Parkinson's disease patients? *Parkinsonism Relat Disord* Date of Electronic Publication: 2019 DOI: [10.1016/j.parkreldis.2019.02.019](https://doi.org/10.1016/j.parkreldis.2019.02.019)
- Savitz, J., y Jansen, P. (2003). The stroop color-word interference test as an indicator of ADHD in poor readers. *The Journal of Genetic Psychology*, 164(3) 319-333.

- Schependorn, J., D'hooghe, M., Cleynhens, C., D'hooghe, M., Haelewyck, M., Keyser, J., y Nagels, G. (2015). Reduced information processing speed as primum movens for cognitive decline in MS. *Multiple Sclerosis Journal*, 21(1) 83–91
- Scheres, A., Oosterlaan, J., Geurts, H., Morein-Zamir, S., Meiran, N., Schut, H., Vlasveld, L., y Sergeant, J. (2004). Executive functioning in boys with ADHD: primarily an inhibition deficit? *Archives of Clinical Neuropsychology* 19, 569–594
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6), 323-338.
- Scretlen, D. J. (2010). Modified Wisconsin Card Sorting Test: M-WCST; Professional Manual. PAR
- Settle, J., Robinson, S., Kane, R., Maloni, H., y Wallin, M. (2015). Remote cognitive assessments for patients with multiple sclerosis: a feasibility study. *Multiple Sclerosis Journal*, 21(8) 1072–1079
- Sherer, M., Nick, T., Millis, S., y Novack, T. (2003). Use of WCST and the WCST-64 in the assessment of traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 25(4), 512-520.
- Shum, D., McFarland, K., y Bain, J. (1990). Construct validity of eight tests of attention: Comparison of normal and closed head injured samples. *Clinical Neuropsychology*, 4(2), 151-62.
- Siegrist, M. (1997). Test-retest reliability of different versions of the Stroop test. *The Journal of Psychology*, 131(3), 299-306.
- Smith, A. (1968). The symbol-digit modalities test: A neuropsychologic test of learning and other cerebral disorders. *Learning Disorders*. Edited by: Helmuth J.

- Smith, A. (2002). *Manual de test de símbolos y dígitos SDMT*. Publicaciones de psicología aplicada. TEA ediciones.
- Solhberg, M., y Mateer, C. (2001). *Cognitive Rehabilitation. An integrative neuropsychological approach*. The Guilford Press.
- Soprano, A. (2003) Evaluación de las funciones ejecutivas en el niño. *Revista de Neurología*, 37 (1) : 44-50.
- Soto, H., Cairo, E., Marrero, M., y González, E. (2016). Síndrome disejecutivo en pacientes con lesiones cerebrales no frontales. Evidencias empíricas obtenidas con la aplicación del WCST. Cuadernos de Neuropsicología/Panamerican. *Journal of Neuropsychology*, 10(2), 55-70.
- Strauss, E., Sherman, E., y Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. 3th Edition*. Oxford University Press.
- Strickland, T. L., D'elia, L. F., James, R., y Stein, R. (1997). Stroop color-word performance of African Americans. *The Clinical Neuropsychologist*, 11(1), 87-90
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662
- Stuss, D., Levine, B., Alexander, M., Hong, J., Palombo, C., Hamer, L., Murphy, K., y Izukawa, D. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia* 38, 388-402
- Takeda, N., Terada, S., Sato, S., Honda, H., Yoshida, H., Kishimoto, Y... Kuroda, S. (2010). Wisconsin Card Sorting Test and Brain Perfusion Imaging in Early Dementia *Dement & Geriatr Cogn Disord* 29, 21–27.

- Ternes, A. M., Clough, M., Foletta, P., White, O., y Fielding, J. J. (2019). Clin Exp. *Neuropsychol.* 20:1-7. doi:10.1080/13803395.2019.1614536
- Ter-Stepaniem, M., Grizenko, N., Cornish, K., Talwar, V., Mbekou, V., Schmitz, N., y Joobar, R. (2017). Attention and Executive Function in Children Diagnosed with Attention Deficit Hyperactivity Disorder and Comorbid Disorders. *J Can Acad Child Adolesc Psychiatry*, 26:1, 21-30.
- Unnanuntana, A., Jarusriwanna, A., y Nepal, S. (2018). Validity and responsiveness of Barthel Index for measuring functional recovery after hemiarthroplasty for femoral neck fracture. *Archives of Orthopaedic and Trauma Surgery*, 138, 1671-1677
- Urtasun, M., Daray, F., Teti, L., Coppolillo, F., Herlax, G., Saba, G...Irazola, V. (2019). Validation and calibration of the patient health questionnaire (PHQ-9) in Argentina. *BCM Psychiatry*, 19: 291 <https://doi.org/10.1186/s12888-019-2262-9>
- Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., y Jolles, J. (2006). The stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment*, 13(1), 62-79.
- Van der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., y Jolles, J. (2008). Detecting the significance in performance on the Stroop Color-Word Test, Rey's Verbal Learning test, and the Letter Digit Substitution Test: The regression -based model app *Journal of the International Neuropsychological Society*, 14, 71-80.
- Vanotti, S., Cores, E., Eizaguirre, B., Angeles, M., Rey, R., Villa, A., y Caceres, F. (2015). Normatization of the Symbol Digit Modalities Test-Oral Version in a Latin American Country *Applied Neuropsychology: Adult*, 22, 46-53



- Vaurio, L., Riley, E., y Mattson, S. (2008). Differences in executive functioning in children with heavy prenatal alcohol exposure or attention-deficit-hyperactivity disorder *Journal of the International Neuropsychological Society*, 14, 119–129.
- Valdez, P., Nava, G., Tirado, H., Frías, M. y Corral, V. (2005). Importancia de las funciones ejecutivas en el comportamiento humano: implicaciones en la investigación con niños. En M. Frías y V. Corral (Dirs.), *Niñez, adolescencia y problemas sociales* (pags. 65-81). México: CONACYT-UniSon.
- Vega, C., y Fernández, M. (2011). Errors on the WCST Correlate with Language Proficiency Scores in Spanish-English Bilingual Children *Archives of Clinical Neuropsychology*, 26, 158–164
- Villamizar, G., y Guevara, E. (2013). Bilingüismo y desempeño de las funciones ejecutivas en adolescentes de colegios bilingüe y no bilingüe. *Revista de Pedagogía*, 34 (94), 35-58
- Villaseñor-Cabrera, T., Guardia-Olmos, J., Jiménez-Maldonado, M., Rizo-Curiel, G., y Peró-Cebollero, M. (2010). Sensitivity and specificity of the Mini-Mental State Examination in the Mexican population. *Quality and Quantity*, 44, 1105-1112.
- Vitkovitch, M., Bishop, S., Dancey, C., y Richards, A. (2002). Stroop interference and negative priming in patients with multiple sclerosis *Neuropsychologia* 40, 1570–1576
- Vogel, A., Stokholm, J., y Jørgensen, K. (2013). Performances on Symbol Digit Modalities Test, Color Trails Test, and modified Stroop test in a healthy, elderly Danish sample. *Aging, Neuropsychology, and Cognition*, 20(3), 370-382
- Wang, Q., Sun, J., Ma, X., Wang, Y., Yao, J., Deng, W... y Li, T. (2011). Normative data on a battery of neuropsychological tests in the Han Chinese population. *Journal of Neuropsychology*, 5(1), 126-142

- Yalcinli, S., Ersel, M., Karbek-Akarca, F., Can, O. y Midik, S. (2015). Can Barthel Index predict mortality in geriatric patirnts admitted to the emergency department with a high fever? *Turkish Journal of Geriatrics*, 18 (4), 266-272
- Zhang, L., Ding, C., Li, H., Zhang, Q., y Chen, A. (2013). The influence of attentional control on stimulus processing is category specific in Stroop tasks. *Psychological Research* 77, 599–610
- Zimmerman, N., Cardoso, C., Trentini, C., Grassi-Oliveira, R., y Fonseca, R. (2015). Brazilian preliminary norms and investigation of age and education effects on the Modified Wisconsin Card Sorting Test, Stroop Color and Word test and Digit Span test in adults. *Dement Neuropsychology*, 9(2), 120-127